PRODUCTION FORECAST MODELS FOR RENEWABLE ENERGIES

by
Tulası Mohan Telaga

DEPARTMENT OF ELECTRICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY KANPUR
February 2000

Production Forecast Models
for Renewable Energies

A thesis submitted in partial fulfillment
of the requirements for the degree of

Master of Technology

in

Electrical Power Systems

by

Tulasi Mohan Telaga

1017076/9810455

under the guidance of

Prof Dr -Ing Jurgen Stenzel   **Prof Prem Kumar Kalra**

Institut fur Elektrische Energieversorgung   Department of Electrical Engineering
Technische Universitat Darmstadt   Indian Institute of Technology, Kanpur
Germany   India

February 21, 2000

CERTIFICATE

This is to certify that the thesis entitled **"PRODUCTION FORECAST MODELS FOR RENEWABLE ENERGIES** submitted by **Tulasi Mohan Telaga** to Indian Institute of Technology Kanpur for the award of Master of Technology in Electrical Engineering is a bonafide record of project work carried out under our supervision He is jointly guided by **Prof Dr Ing Jurgen Stenzel** Technische Universitat Darmstadt Germany and **Prof Prem Kumar Kalra** Indian Institute of Technology Kanpur India Contents of the thesis in full or in parts have not been submitted to any other institute or university for the award of degree or diploma

Dr Ing Jurgen Stenzel

Professor

Institu for Electrische Energieversorgung

Technnische Universitat

Darmstadt Germany

Dr Prem Kumar kalra

Professor

Department of Electrical Engg

Indian Institute of Technology

Kanpur India

INSTITUT FUR ELEKTRISCHE ENERGIEVERSORGUNG

TECHNISCHE UNIVERSITAT DARMSTADT

Prof Dr Ing J Stenzel

Certification

Telaga, Tulasi Mohan * 7 4 1977

Indian Institute of Technology Kanpur India (Prof Prem K Kalra)

Evaluation of the Master Thesis

Production Forecast Models for Renewable Energies

Under my supervision and the guidance of Dipl -Ing Kristjan Halldorsson Mr Telaga prepared his thesis After being introduced into the subject he worked independently and he was very diligent Despite some problems in finding relevant data for testing his models his research work was very successful The presentation and discussion of his thesis testified his good understanding of the problem Considering the various aspects of his thesis (research work results written thesis and presentation) my evaluation leads to the mark **very good**

(German marks 1 very good 2 good 3 satisfactory 4 sufficient 5 unsatisfactory)

During his stay in Germany Mr Telaga had the opportunity to spend some time in industry

- **repas AEG Automation GmbH** Dreieich (2 weeks)
- **HEAG**, Darmstadt (local utility   2 weeks)

and to join an

- **Excursion** to Austria (Industry and Power Plants   week)

organised by the Department of Electrical Engineering TU Darmstadt

Darmstadt 20 2 2000

(Prof Dr Ing Jurgen Stenzel)

Acknowledgements

Abstract

In the last years many countries in the world have opened their power markets to allow competition between power producers At the first look it seems that the renewable energies do not have a chance in those deregulated markets because they are more expensive than the conventional one and some of them are having problems because of fluctuations in the production In this thesis the fluctuation problem is addressed and forecast models for the production of renewable energies mainly wind and photovoltaic are developed This will help the trading of these renewable energies in deregulated power market For forecasting the production of renewable energies fuzzy neural network and regression models are utilized and tested with the actual measured data

Contents

1. Introduction

In the last years many countries in the world have opened their power markets, to allow competition between power producers At the first look, it seems that the renewable energies do not have a chance in those deregulated markets, because they are more expensive than the conventional one and some of them are having problems because of fluctuations in the production

The main advantage of the renewable energies compared with the conventional energies is that they are not causing any appreciable environmental disturbances, like CO_2 increase and greenhouse effect Because of these advantages the pro duction of renewable energies is subsidized in some of the industrial countries, which already are deregulated The purpose of these subventions, is to make the renewable energies competitive at the power market

Although the renewable energies were made competitive, their position in the deregulated market would still not be simple, because of fluctuations in the pro duction This applies specially to wind and photovoltaic energy To make the trading of those "fluctuating renewable energies" possible, forecasts for their pro- duction is obviously of great importance With such a forecast, the buyers of renewable energies, have the chance to compensate the coming fluctuations in the renewable production through trading at some conventional power markets, for example the spot market

In this work some forecast models for renewable energies will be designed, im plemented and compered The models will be limited to wind and photovoltaic energy that is the most fluctuating ones For the wind power forecast a model based on fuzzy logic will be compered with a regression analysis model For the photovoltaic power forecast a neural network model will be established and also

compered with regression analysis model

In next chapter the trading of renewable energy in deregulated market and the need for power forecasts will be discussed further   Chapter 3 and 4 handle the main theme of this work, the wind and the photovoltaic power forecasting models   Chapter 5 and 6 deal with the basics for multiple regression and neuronal networks, that is theory that is used for the wind and the photovoltaic models   The conclusions of the work are finally in chapter 7

2 The Trading of Renewable Energy in a Deregulated Market

In this chapter the problematic concerning the trading of renewable energy in a deregulated market will be discussed. The two main problems that a renewable energy marketer has to deal with are

- The expensiveness of the renewable energy

- The fluctuations of the renewable production

In practice there already exist some techniques to deal with the first problem, for example the renewable order in England and Wales [17] and the green label approach in Holland [22]. Those techniques have in common that their aim is to make the renewable energy competitive at the power market through some kind of furtherance. The second problem, the fluctuations of the renewable production, is harder to handle. A power marketer, who is purchasing or selling renewable energy, has to take those fluctuations into account.

Here the focus is on the second problem. For that purpose a simple model of a deregulated power market will be set up and the situation of a power marketer who is trading with renewable energy, will be analysed. It is admitted that the renewable production is promoted in some way, so it is competitive. It will be shown that the marketer will definitely need a forecast for the production of the renewable energy.
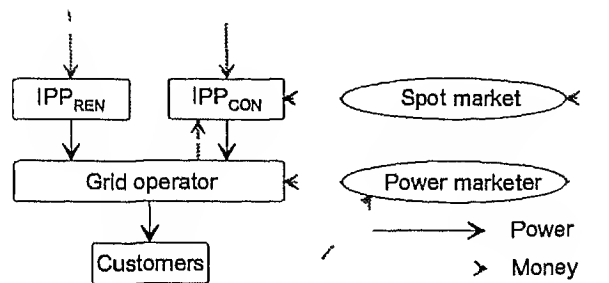
Figure 2 1   A model of a deregulated power market

2 1   A Model for a Deregulated Power Market

Figure 2 1 shows a simple model of a deregulated power market   The market is strictly divided into 4 sectors   Production, transport, marketing and distribution The power production is in the hands of *independent power producers* ($IPP's$) They are classified as *renewable* and *conventional*   A *grid operator* is responsible for the transport of power from the $IPP's$ to the customers   The grid operator does not produce any power at his own   He will buy regulation power directly from some conventional power producers   The *spot market* is an organised day ahead market for electric energy [20]   The participants at the spot market give their offer how much they want to buy/sell and for what price the day before the activation of trade   A *power marketer* buys power from the $IPP's$ and sell it to his customers   The customer do not handle directly with the $IPP's$ It is assumed that the analytical method [5] is used, that means that the power marketer has to pay the difference between the consumption of his customer and the estimated demand   A power marketer has 4 ways to buy energy for his customers, as indicated in figure 2 2   Those 4 ways are

1   Buying energy through bilateral trading from conventional IPP

2   Buying energy through bilateral trading from renewable IPP

3   Buying energy through the spot market

4   Buying regulation power from the grid operator
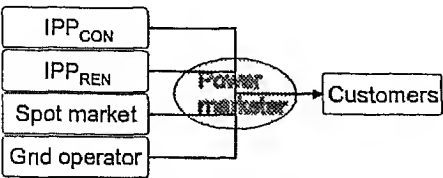
Figure 2 2   The possible ways for a power marketer to buy energy for his cus
            tomers

If the power marketers haven't order enough power through bilateral trading or
through the spot market to fulfil the consumption of their customers , the grid
operator has to buy regulation power at the regulation market to balance out the
difference between consumption and production   The power marketers have to
pay for the regulation power

2 2   Trading of Renewable Energy

Now the situation of a power marketer, who is trading with renewable energy,
will be considered   It is assumed, that the marketer is having contract with
one or more $IPP's_{REN}$ and that he is buying all the power that they can
possibly produce   Moreover it is admitted that the power marketer is buying
some constant amount of power from one or more $IPP's_{CON}$ He wants to buy
the rest, that he needs to fulfil his customers consumption, from the spot market
Figure 2 3 shows this for a one particular day   The power marketer has to decide
the day before this particular day, how much he wants to order from the spot
market for each period of the following day   Here each day is divided into 24
periods, that is, each period is 1 hour   What the power marketer needs from the
spot market in some particular period is, as indicated in figure 2 3, the estimated
consumption minus what he will get from the $IPP's$ It's clearly important that
the marketer has some good forecasts for the consumption of his customers and
the production of the $IPP'_{REN}s$ for the following day   Here (in Figure 2 3) is
admitted that he has such forecasts, with constant values during each period
Figure 2 4 shows the difference between the estimated and real consumption for
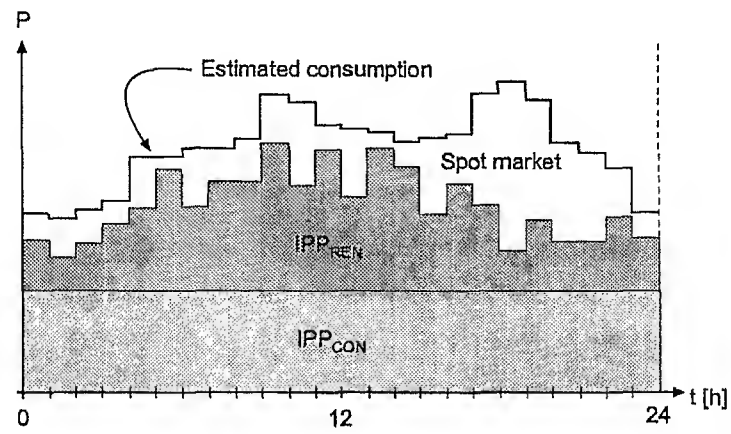
Figure 2.3.: The purchase of power of a power marketer in on day
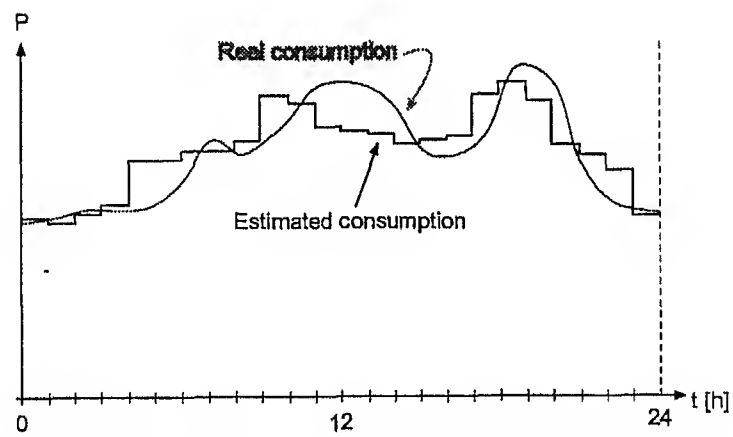
Figure 2.4.: Estimated and real consumption

this one particular day The grid operator balance out the differences (positive or negative) and the power marketers have to pay for it afterwards So good estimation for the consumption is not just helpful, when ordering from the spot market, but also extremely important because of the regulation costs

2 3 Needed Forecasts for a Power Marketer

As a result of this chapter the following can be stated A power marketer who is responsible for selling power from some $IPP's_{REN}$, and who wants to compensate the fluctuations of the renewable production with purchase at the spot market, needs the following forecasts for the coming day

- The estimated consumption for each period

- The estimated production of the $IPP's_{REN}$ for each day

In this report the second problem, that is the estimated production of the $IPP's_{REN}$ for each day, will be considered This work is concentrated on wind and photo valtiac power forecasts, because those are the renewable energies were one can expect the most fluctuations

3 Wind Power Forecasting

3 1   Introduction

Wind energy converters (WEC) integrated in electrical power systems may cause several problems, including power quality, stability and especially power dispatch ing   These problems increase as the penetration of wind energy increases   The power generated by wind turbines changes rapidly because of the continuous fluctuation of wind speed and direction   As explained in previous chapter it is extremely important, for the trading of wind energy in deregulated power market, to have good forecast for the wind power production at any specified time for the following day

In the literature, various models, have been developed for wind speed or power prediction   The time scale of these models ranges from some seconds, when the objective is wind turbines control, to some minutes, or even hours, when the ob jective is economic dispatch and powersystem planning   The time series analysis is common model for small time scales   In higher time scales, of some hours or more, models incorporating meteorological information are usually used   The time series models found in the literature do not provide a significant improvement over the method of persistence   As a consequence, the latest one, being extremely simple (the wind in all future time steps will be equal to the wind now), is almost exclusively used in the applications [15]   In [6] ARMA (auto regressive moving average) models for wind speed prediction achieve an improvement with respect to the persistance model   In [4] multilayer feed-forward neural networks and radial basis functions are applied for wind power prediction   In [9] various models like ARMA and bilinear ones are developed for wind power   In

[1] an artificial neural network based model which uses local and spatial relations of the wind speed is proposed for estimating wind power

In this chapter probability and fuzzy based model is developed for short term prediction of wind power   This model uses wind velocity and its regularity to forecast the wind power   The basic idea is developing a possibility curve for wind power generation at a forecasted wind speed and apply some fuzzy operations, like fuzzification, composition and defuzzification, to get an approximate wind power   Wind power possibility curve is obtained by manipulating the historical data   At last regression analysis is also used to make comparison for the proposed model

This chapter is divided into 5 sections   In section 3 2 the basics of the forecasting methodologies are explained   Section 3 3 explains the detailed description of the proposed model where as section 3 4 discusses about the regression model   In section 3 5 results are presented for the proposed fuzzy model as well as regression model and compared both   In section 3 6 conclusions regarding the proposed model with respect to the regression model are discussed

3 2   The Basics of Wind Power Forecasting

The power generated by a wind turbine is chiefly influenced by wind velocity and direction [26]   However, many other factors such as the turbines immediate history of power generation, the blade position at the time of measurements [21] dynamic load distribution among parallel turbines, alternation of system load flow, dynamic performance of a generator, and wind dynamics also influence gen erated power, so that even when the average wind speed within 10 minutes is the same, the power generated can be very different [24]

The direction of wind also influence power generation   However, compared with wind velocity, wind direction has less influence on power output [24] because most of the turbines are built to turn into the wind when operating   Generally, at the same level of wind speed, there is no great difference of the power generation for

different wind direction

There are two main approaches that may be followed in order to generate wind power forecast

1 to develop an explicit prediction model for wind power, in which it will be possible to consider wind speed, wind direction etc   as explanatory variables

2 to develop a prediction model for wind speed and a second model for the transformation of wind speed to power

In this project the first methodology is used for forecasting wind power in which wind speed and wind speed regularity are considered as input parameters   The detailed model is discussed in next section   Here are presented Naive predictors for the wind power forecasting in which only data manipulation is involved

3 2 1   The Naive Predictors

When an advanced forecasting model is not available, forecast may be obtained with a minimal effort and data manipulation and can be based solely on the most recent information available   Such forecasts are referred to as Naive forecasts One such method (Persistent or Naive 1 method), is to use the most recent datum available P(t), as forecast $P(t + k/t)$ for each one of the future time steps that is P(t+k/t) = P(t) for $k = 1, 2,$ n A slightly more sophisticated method would be to use the average of m past values (Naive m) as forecast

$$P_m(t + k/t) = \frac{1}{m} \sum_{i=0}^{m-1} P(t - i), k = 1, 2, \quad n \tag{3 1}$$
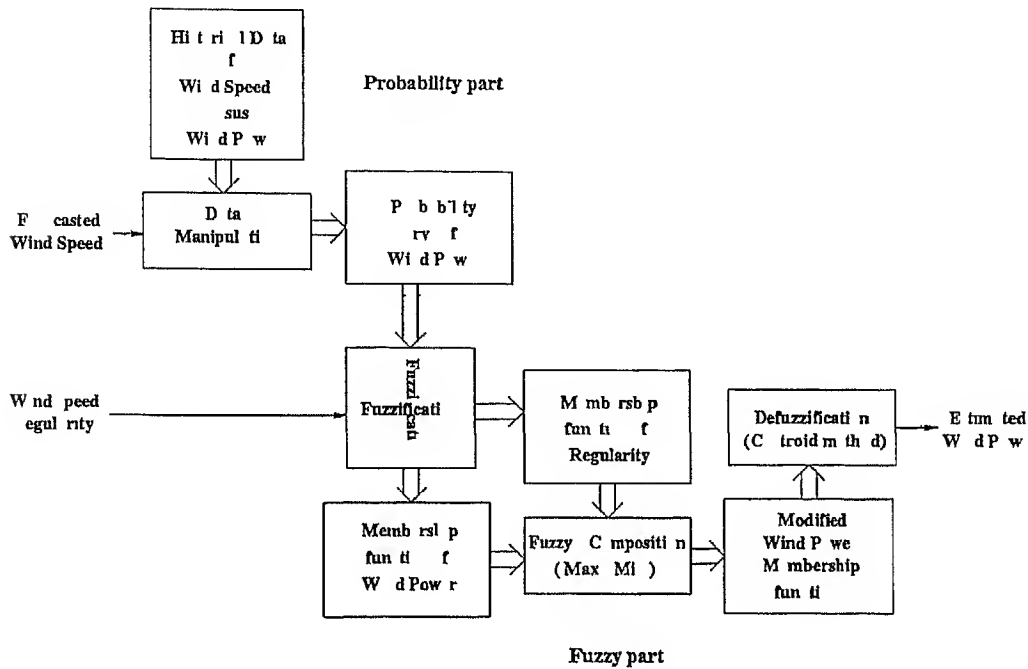
Figure 3 1   Fuzzy based model

3 3   The Fuzzy Based Model

This model can be divided into two parts   One is the probability part and the other is the fuzzy part as shown in figure 3 1   In the probability part a probability curve of the wind power generation at a forecasted wind velocity will be developed by using historical data   In the fuzzy part the probability curve will be manipulated through some fuzzy operations according to information about the regularity

3 3 1   The Probability Part

In this section developing probability curve of the wind power generation at a forecasted wind speed is processed   This is done by using the historical measurements that are taken for wind speed and power generation   As shown in figure 3 1 the probability part is having two inputs   One is *historical data* and the other one is *forecasted wind speed*   Output of this block is *the probability curve*

of the wind power In the data manipulation block, some statistical analysis of historical data with respect to the forecasted wind speed will be carried Now the data manipulation process will be explained It is admitted that x is wind speed and y the corresponding wind power

Consider there are N pairs of measurements $(x_1, y_1), (x_2, y_2),$, (x_n, y_n) between two variables x and y are in historical data The first step in data ma nipulation block is analyzation of the relation ship between x and y As indicated in figure 3 2, there may be several values of wind power for one value of wind speed As mentioned before that the output of the data manipulation process is a probability function for the wind power This probability function indicates the different values of wind power that one could expect for the forecasted wind speed according to the historical data Now the composition of this probability function will be be discussed

For the composition of the probability curve only the points in the historical data, which lay between $x_o - \Delta x$ and $x_o + \Delta x$, will be used (see figure 3 2) Those points are $(x_1, y_1), (x_2, y_2),$, (x_M, y_M), where M is the number of points laying in this range Here x_o is the forecasted wind speed and Δx is the clustering range around the x_o

Selection of the size of Δx depends on the density of the data points around x_o If the density is high, Δx should be small and vice versa This is reflected in the below two rules for Δx

1 Δx should not exceed certain range that is Δx_{max}, and

2 The number of points (M) should not exceed certain value that is M_{max}

Δx_{max} and M_{max} are constants, which the user can vary depending on his experience with the model

After getting the points, it is better to do normalization to the above points, by using below equation

$$x_{i(norm)} = \frac{x_{i(original)}}{max(x_1, x_2, \quad , x_M)} \times 100 \qquad (3\ 2)$$
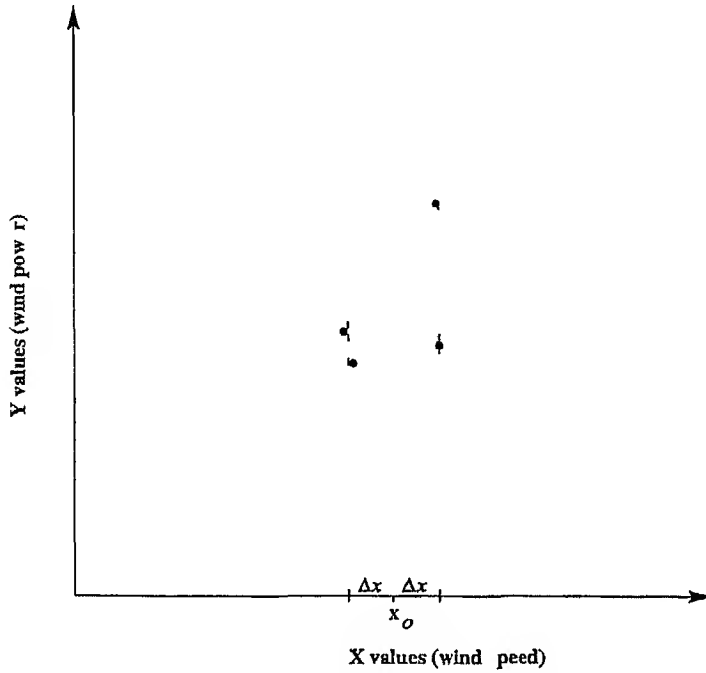
Figure 3 2   Wind speed versus wind power graph

Here $x_{i(norm)}$ is the normalized value and $x_{i(original)}$ is the original value   This normalization is for making the computation easy

The probability function is nothing but the comparison among the points, so here comparison between the above points is done with it's mean, which can be calculated by using below equation

$$x_c = \frac{x_{1(norm)} + x_{2(norm)} + \quad + x_{M(norm)}}{M} \qquad (3\ 3)$$

$$y_c = \frac{y_{1(norm)} + y_{2(norm)} + \quad + y_{M(norm)}}{M} \qquad (3\ 4)$$

Here x_c and y_c represents the mean point among the M points

The distances of these points from its mean is the one way to compare them Here the distances of each point from its mean is calculated and is denoted with d_i in below equation

$$d_i = \sqrt{(x_c - x_{i(norm)})^2 + (y_c - y_{i(norm)})^2} \qquad (3\ 5)$$
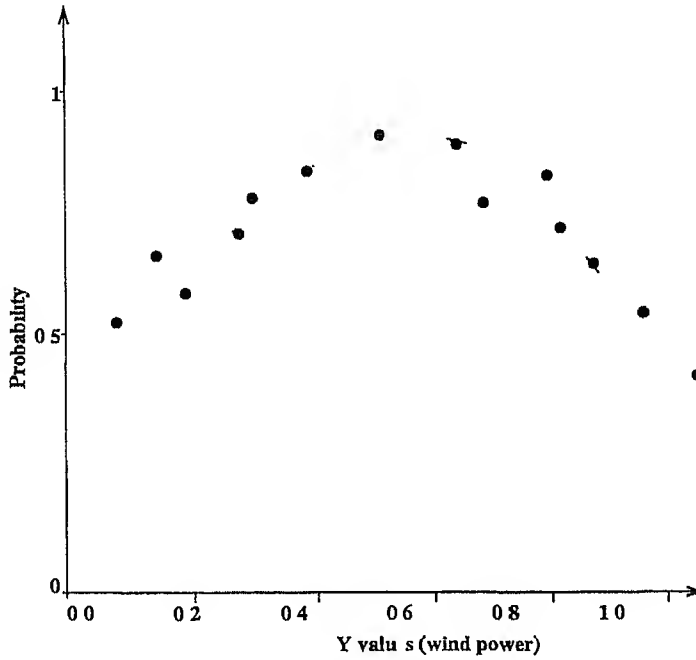
Figure 3 3   Probability of wind power at forecasted wind speed

One can observe from the figure 3 2 that for the different wind speeds, there may be the same wind power It means that, in the above points there may be same value of y for different x values Here the process is to find the probability of wind power that means probability of getting *each different wind power* for a forecasted wind speed If there is some points which having same wind power, it could be better to find the average distance of those points, which in turn reflects the average variation of this wind power to the mean wind power Consider there are L different wind power values, in those M points The average distance of each different wind power can be obtained by below formula

$$d_l = \frac{1}{K} \sum_{l=1}^{K} d_k \tag{3 6}$$

where d_l is the average distance of l th wind power and K is the same value of wind power points Now the probability of each wind power (y_l), that is $p(y_l)$ can find by using below equation

$$p(y_l) = \frac{100 - d_l}{100} \tag{3 7}$$

After finding all the probability values for different wind power values, one can get the probability curve of the wind power generation at a forecasted wind speed as shown in figure 3 3

3 3 2   The Fuzzy Part

This fuzzy part is having two inputs, one is the above probability function of the wind power at a forecasted wind speed and wind speed regularity informa tion The output of this fuzzy part is the estimated wind power for a forecasted wind speed Wind speed regularity information tells about the accuracy of the forecasted wind speed In this fuzzy part fuzzy operations like fuzzification of input variables, fuzzy composition and defuzzification is carried, which will be explained now

3 3 2 1   The Fuzzification Process

Fuzzification is carried for the input variables, that is "probability function of wind power generation" and "wind speed regularity information" Fuzzifica tion means converting the input variables into fuzzy representative membership function format Here different fuzzification processes are applied for the input variables Next in this section fuzzification process for each input variable is ex plained

Fuzzification of wind speed regularity

Regularity itself is a fuzzy term, so one can take wind speed regularity infor mation as a fuzzy representation of natural language qualifiers "Very Regular", "Regular",and "Little Regular" displayed in figure 3 4 It can be represented in a functional form as shown below [18]

$$f(x) = 1 - (1 - x^{\phi})^{1/\phi} \qquad (3\ 8)$$

The values of ϕ are presented in Table 3 1

Fuzzification of wind power probability curve
Fuzzification of wind power probability curve is done by using normalization tech niques Here normalization technique that is applied for getting fuzzy wind power
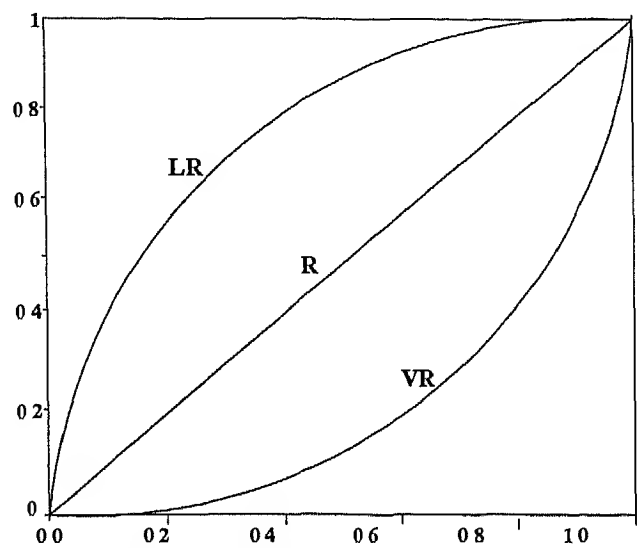
Figure 3 4   Qualifiers for wind regularity

ϕ	Statements
1 5	Very Regular
1	Regular
0 696	Very Little Regular

Table 3 1   Values of ϕ for wind speed regularity

generation is, dividing all the probability values with its maximum probability value It makes the probability curve of the wind power into possibility curve of wind power generation as shown in figure 3 5

3 3 2 2   The Fuzzy Composition

Now there are two membership functions, one is fuzzy wind power generation and other one is wind speed regularity membership function So one can obtain modified fuzzy power generation by applying fuzzy composition rule of inference [19] In [18] traditional modus ponen rules are used to modify the fuzzy power generation

Modus ponens deduction is a very common inference scheme used in forward chaining rule-based expert systems [19] It is an operation whose task is to find the truth value of a consequent in a production rule, given the truth value of
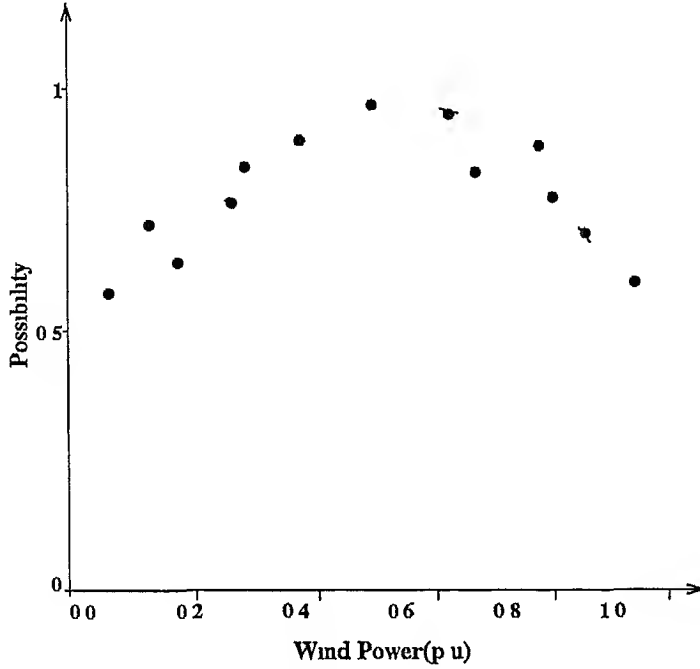
Figure 3 5   Fuzzy membership function of the wind power generation

the antecedent in the rule   Modus ponens deduction concludes that given two propositions P and $P \rightarrow Q$ both of which are true, then the truth of the simple proposition Q is automatically inferred by below equation

$$(P \wedge (P \rightarrow Q)) \longrightarrow Q \qquad (3\ 9)$$

In [19] different types of fuzzy composition rules are discussed in which "max min" composition is mostly used

The final fuzzy power generation can be get by using "max min" composition which is presented below

$$\mu_{PoM}(x, y) = \sup_{y \in Y} min(\mu_P(x, y), \mu_M(y\ z)) \qquad (3\ 10)$$

where x is fuzzy value for power

y is membership value for the basic models, and

z is modified possibility values

When a qualifier such as "little regular" is applied (curve LR of figure 3 4) , low values of possibility rise, and the resulting distribution is more "wide" , meaning
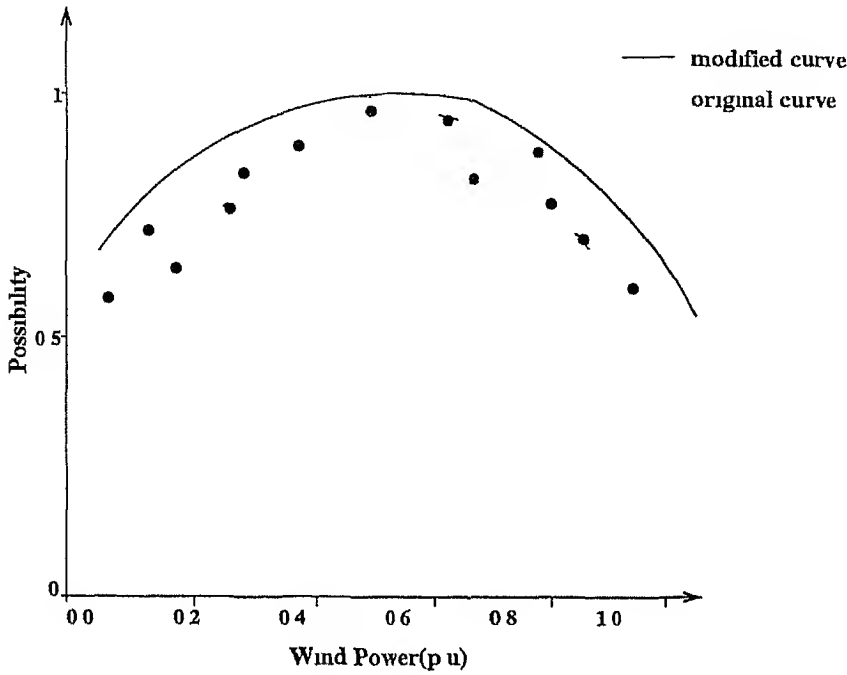
Figure 3 6   Modified wind power membership function

that we are less certain about the true value than in the base case This can be shown graphically as shown in figure 3 6 If the qualification is regular then the original function remains unchanged

3 3 2 3   The Defuzzification Process

After getting the modified fuzzy power generation, one can get the crisp value of power by using the fuzzy defuzzification techniques In liteiature [16] many defuzzification techniques aie discussed, but the "centroid method" is the most populai one

The "centroid method" will now be explained, with the help of figure 3 7   In discrete case where $x = [x_1, x_2, \quad , x_l]$ and corresponding membership values $\mu(x) = [\mu_1(x),_m u_2(x), \quad , \mu_l(x)]$, then by using centroid method the crisp value of x is

$$x^* = \frac{\sum_{i=1}^{l} x_i \mu_i(x)}{\sum_{i=1}^{l} \mu_i(x)} \tag{3 11}$$
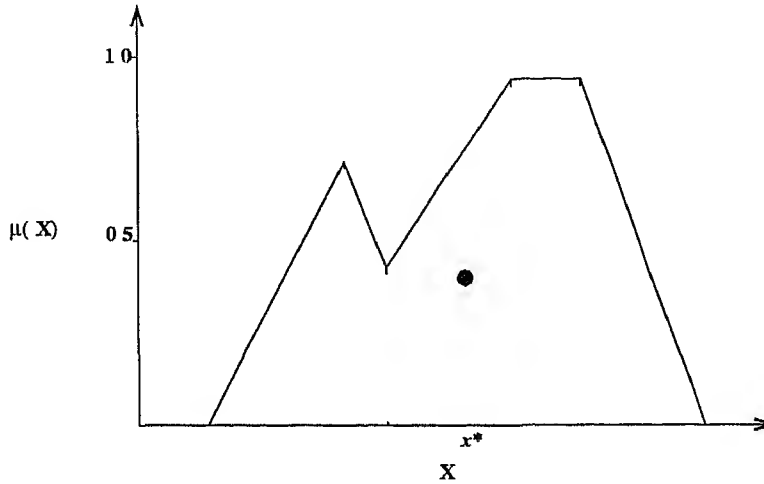
Figure 3 7   Centroid defuzzification method

By employing "Centroid" defuzzification method, for the modified wind power membership function one can get the estimated wind power for the forecasted wind speed

3 4   The Regression Model

Regression based prediction model also can be used for wind power forecasting Regression analysis involves basically statistical analyzation of the historical data One can write the regression equation for wind power forecasting as shown below

$$Windpower = a + b(Windspeed) \qquad (3\ 12)$$

where a and b are the regression equation coefficients, which are to be determined by using historical data   The detailed description of the regression technique is explained in chapter 5

3 5   Results and Comparisons

The data that is taken to test the proposed model is from the "Norden Ostermarch" wind power plant in Germany   They measures the data with a 5 minutes interval in each day and they measures mainly wind velocities and wind power   The

measurement of wind speeds are taken at a hight of 10m from the sea level The rated power of this plant is 500kW, and rotor diameter is 37m which is located 40m high from the sea level

The data taken from the Norden Ostermarsch, for testing the above both models is 5 minutes interval data of wind velocities and wind power generation from 11th October to 14th November 1999 In this data the first one week data that is 11th October to 17th October, is used as a historical data and two other days (8th October and 14th November) are selected for testing the model The analysis of the historical data is shown in figure 3 8 In this figure the first one is wind velocity versus wind power graph and other two are time versus wind speed and time versus wind power graphs The same analysis is also carried for the testing data files and are shown in Appendix, figure A 1 and figure A 2

3 5 1   Forecasting Results by Using the Fuzzy Based Model

The above data is used to test the proposed fuzzy model and the results are shown in figure 3 9 and figure 3 10 corresponding to the 18th October and 14th November data   Although there is no information about regularity, here the testing process is carried for all the three cases of the regularity information that is very regular, regular and little regular The results with regularity information such as very regular and little regular are shown in Appendix, figure A 3 and figure A 4, for 14th November data The figures figure 3 9 and figure 3 10, which are already mentioned are corresponding to the "regular" regularity information In Table 3 2 estimation error of the proposed model is presented with the actual and estimated value of wind power Error can be calculated by using below equation

$$Error = \frac{AC - ES}{AC} \qquad\qquad (3\ 13)$$
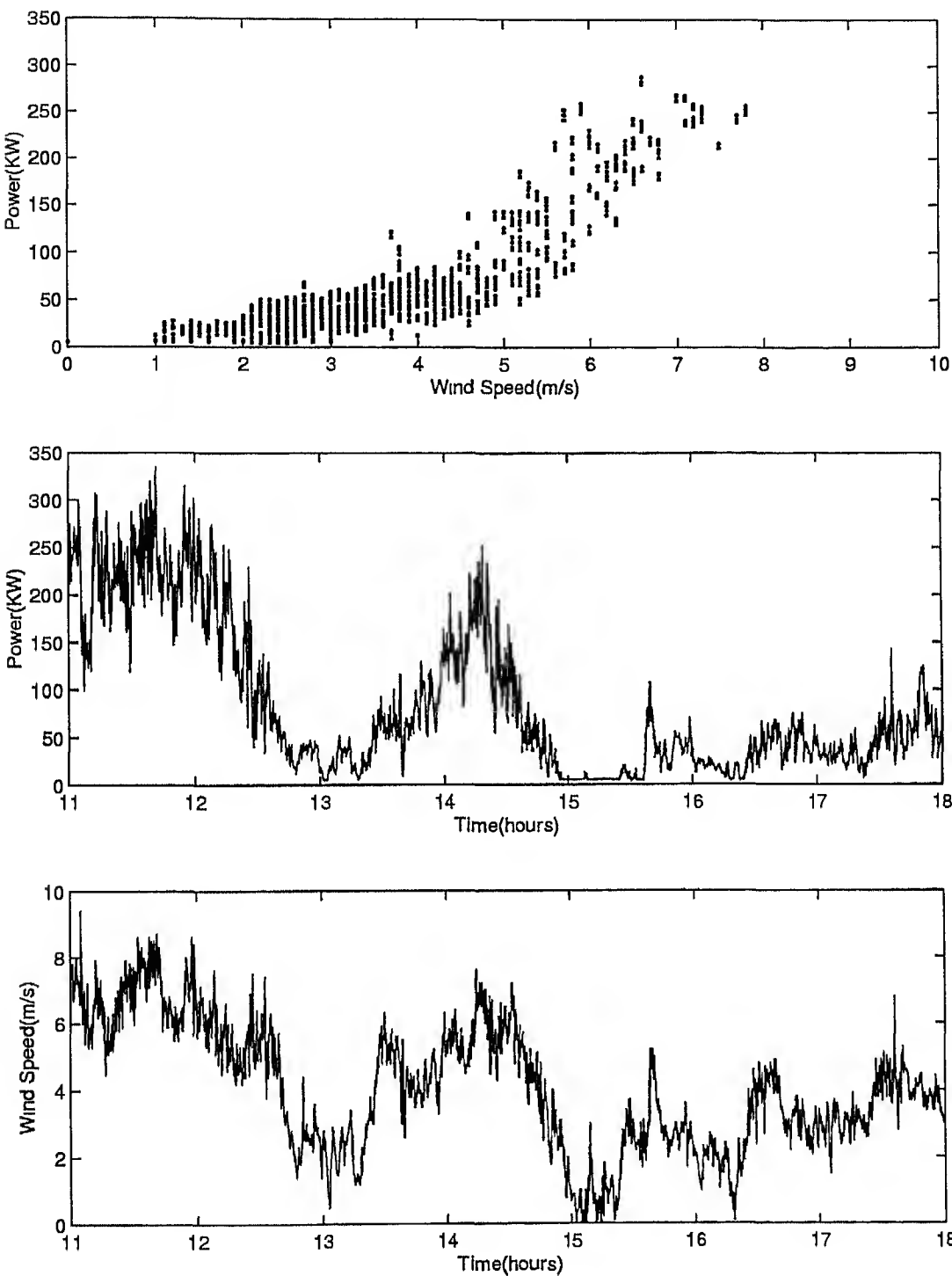
Figure 3 8   11 October to 17 October Oster wind plant data
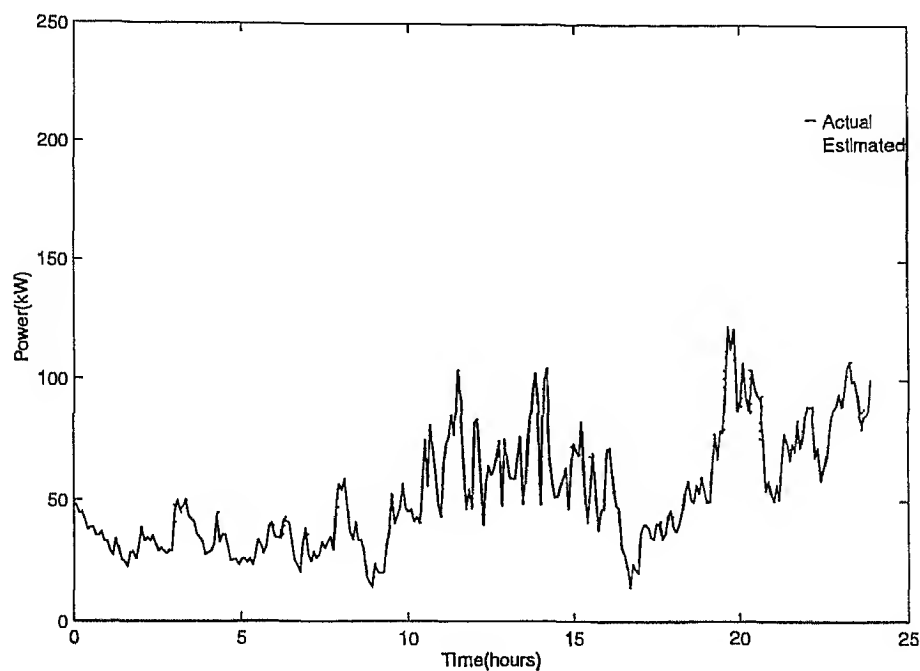
Figure 3 9   18th October Fuzzy based model results
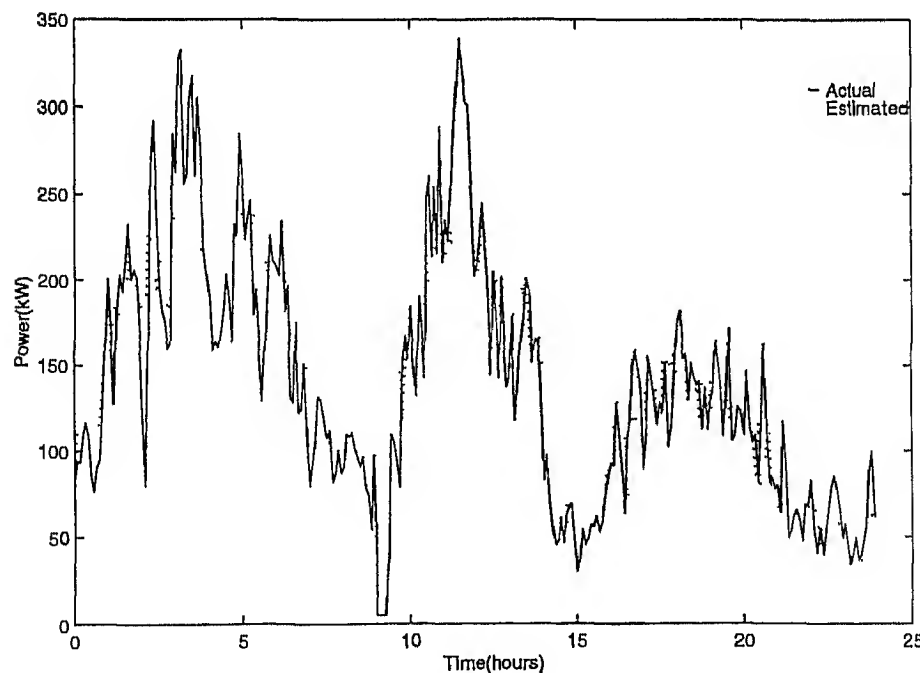
Figure 3 10   14th November Fuzzy based model results

S No	Wind Speed m/s	Wind Speed Regularity	Actual Power kW	Estimated Power kW	Error
1	3 2	V R	48 6	45 75000	0 0586419
2	4 7	L R	63 5	111 417988	0 754613
3	2 7	R	34 4	36 226521	0 0530965
4	5 5	L R	69 3	150 363337	1 169745
5	6 1	R	104 9	175 042783	0 668663
6	4 6	V R	90 2	122 5	-0 3580931
7	6 3	V R	58 5	186 6	2 1897435
8	3 8	L R	52 7	66 506545	0 261983
9	2 4	V R	20 9	36 1	0 727272
10	4 9	L R	124 2	134 715801	0 084668

Table 3 2   Wind power forecasting by using fuzzy model

3 5 2   Forecasting Results by Using the Regression Model

Linear regression analysis is also used in the similar manner as the fuzzy model for forecasting the wind power The data from the 11th October to 17th October is used for the regression analysis and the other two data files are used for testing the regression technique Here wind speed is the only input variable taken into account for estimating the wind power

Testing results of 18th October and 14th November data files are shown in figure 3 11 and figure 3 12 Table 3 3 shows the results of estimated power for specific wind speeds with the Error The complete statistical analysis of the data bank with regression analysis is shown in Table 3 4 In that table the regression coefficients a and b are presented with the multiple regression correlation coefficient R The standard error of b and the hypothesis test results for b and R that is F_b and F_R are also presented in that table F_b and F_R are the F test results, which is a common test in statistics

3 5 3   Comparisons

By observing the fuzzy based model results with the regression based model, it is found that the regression technique is better than the fuzzy model Better results
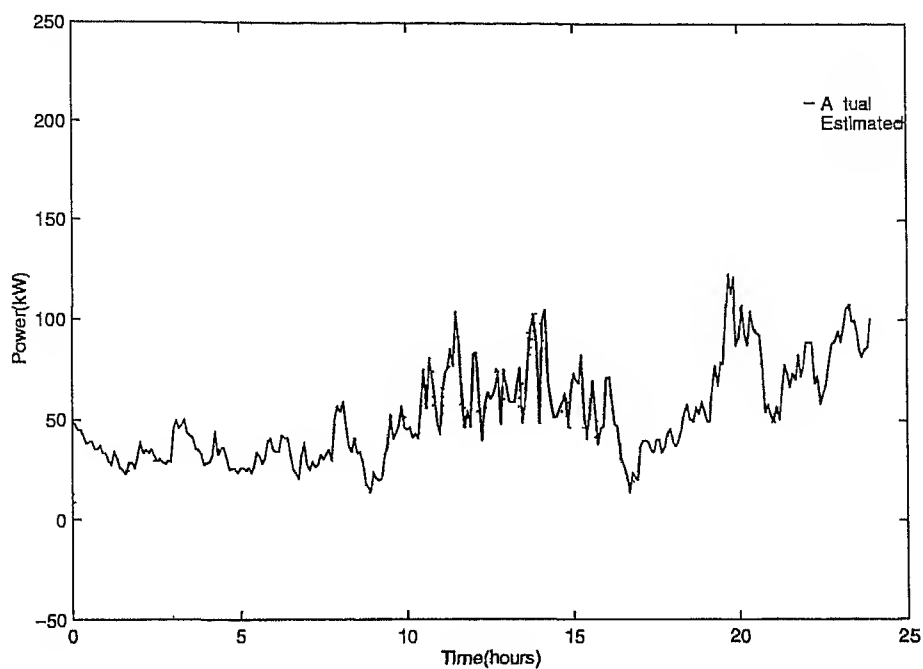
Figure 3 11   18th October multiple regression results
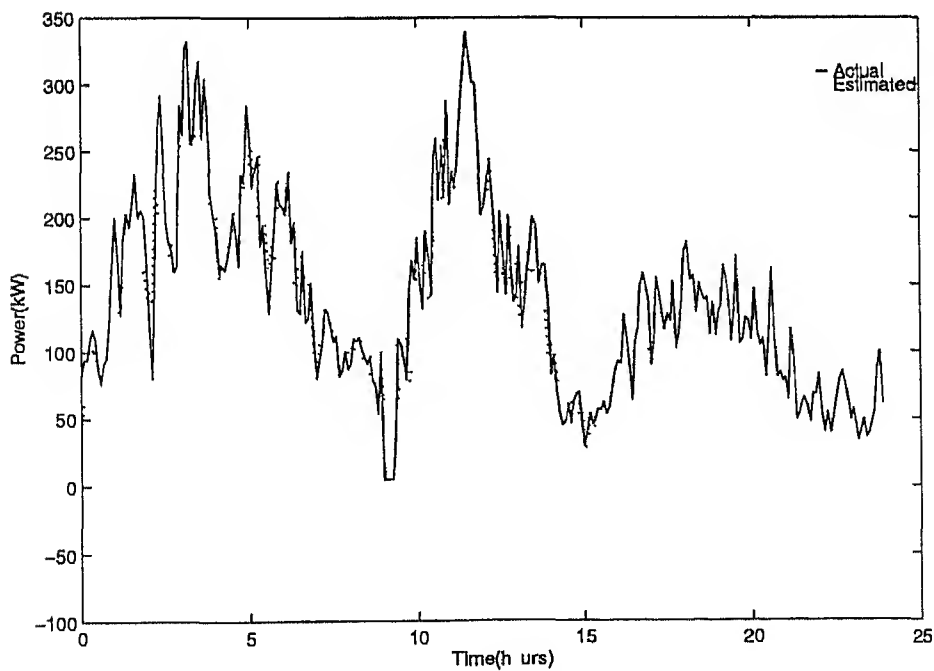
Figure 3 12   14th November multiple regression results

S No	Wind Speed m/s	Actual Power kW	Estimated Power kW	Error
1	3 2	48 6	22 508814	0 53685564
2	4 7	63 5	63 041298	0 007223655
3	2 7	34 4	08 997984	0 73843074
4	5 5	69 3	84 658640	0 22162531
5	6 1	104 9	100 87162	0 03840211
6	4 6	90 2	60 339140	0 33105162
7	6 3	58 5	106 27596	0 8166831
8	3 8	52 7	38 721806	0 26524088
9	2 4	20 9	0 8914852	0 9573452
10	4 9	124 2	68 445640	0 44890788

Table 3 3   Wind power forecasting using multiple regression

Parameter	Regression Coefficient (b)	Standard Error of b	Hypothesis test F_b
Wind Speed	1 0202	0 0134	5827 2
intercept a		-0 1569	
Correlation Coefficient R		0 86206	
Hypothesis test F_R		5827 2	

Table 3 4   Statistical analysis of wind data

obtained from the regression model is can be explained by the statistical Table 3 4 In that the correlation coefficient value R is very close to 1, signifies that the regression equation is best fit for the historical data The standard error of b is also very less The F test results for both regression coefficient and correlation coefficient shows that the both are passed the null hypothesis test The detailed explanation regarding this tests are carried in chapter 5

One main advantage from the fuzzy model is, it can provide the possibility curve of the wind power for the forecasted wind speed, which will provide the better sight about the wind power at that forecasted wind speed The results from the regression model are promissing, but the error is sometimes higher than accessptable value For more better results, one can modify the regression model by changing the regression equation from linear to the nonlinear as given below

$$Windpower = a + b \times Windspeed^3 \qquad (3\ 14)$$

18th October nonlinear regression result $(wind\ power - a + b\ ws^3)$
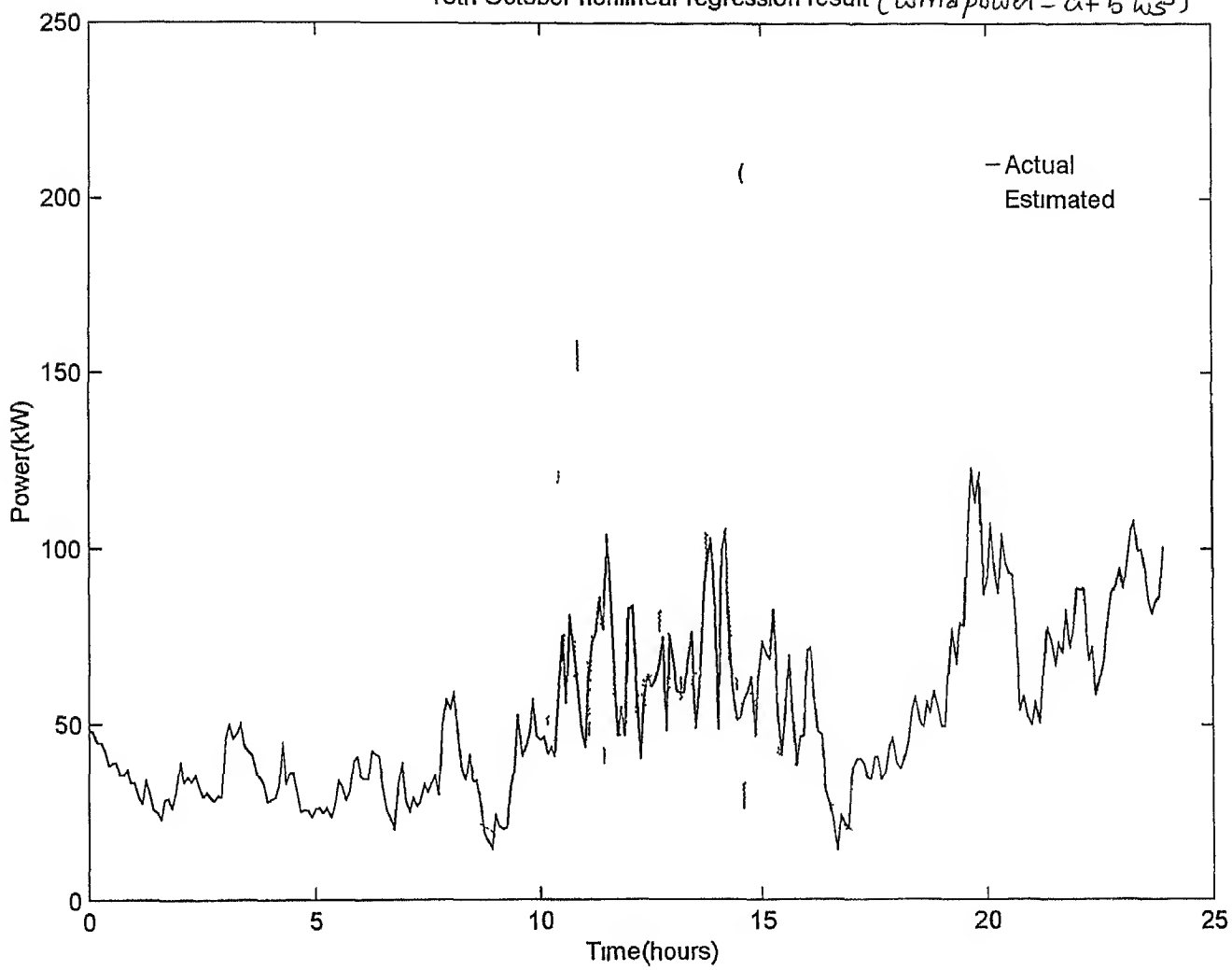
14th November nonlinear regression result $(windpower = a + b \, ws^3)$
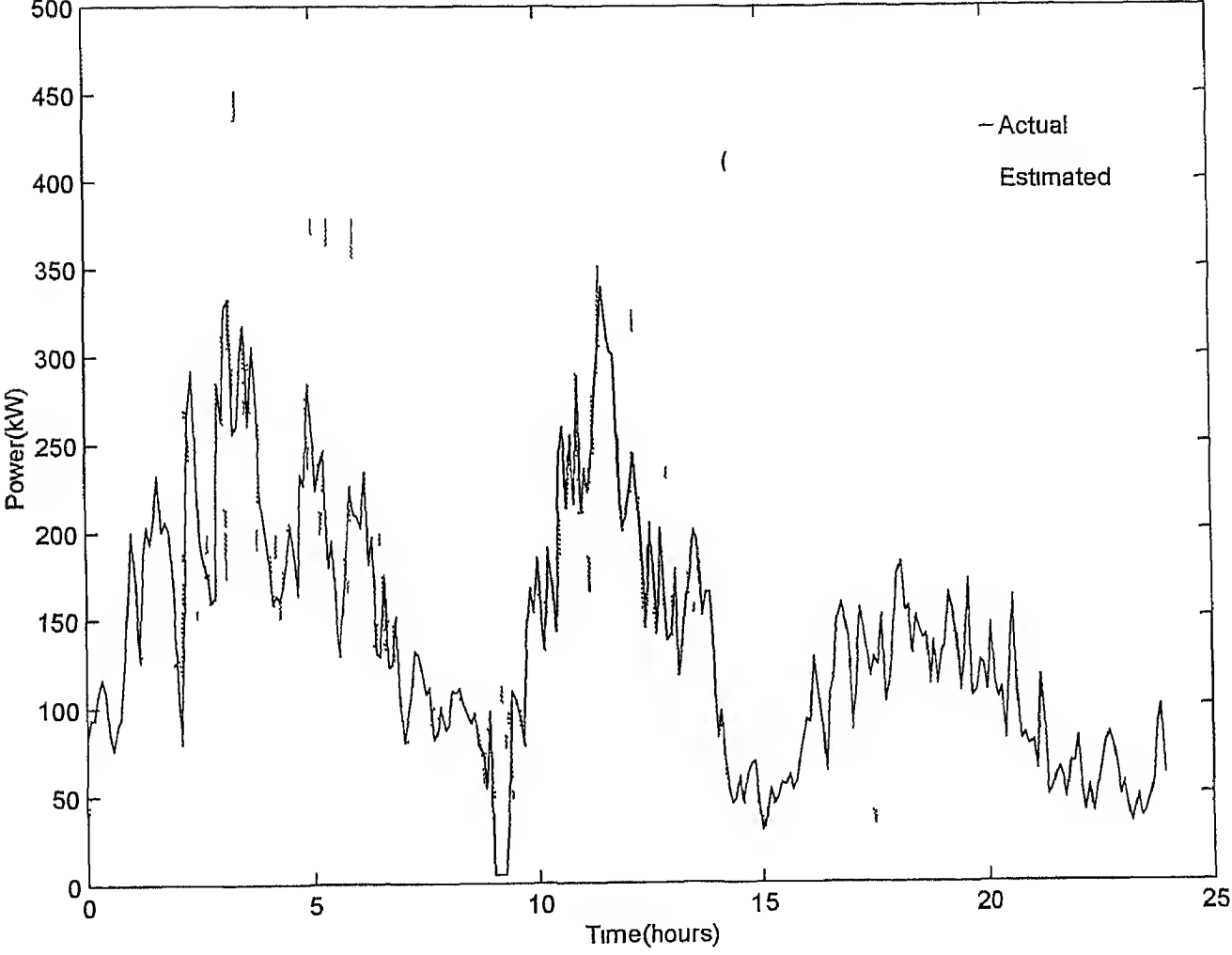
This may give improved results, because it is found in literature that the rela tionship between wind power and wind speed is non linear nature as mentioned above For the proposed fuzzy model one have to go further study on this model to make modifications for getting better results

3 6   Conclusions

In this chapter wind power forecasting models were discussed Two models were constructed for wind power forecasting, a fuzzy based model and a regression based model In fuzzy model, historical data manipulation is carried in the prob ability part, and in fuzzy part, fuzzy operations like fuzzification, composition and defuzzification were used Both the models were tested with the actual wind power plant data From the results it is concluded that the regression based model gave better results than the fuzzy model

4 Photovoltaic Power Forecasting

4 1 Introduction

' Clean electricity with no moving parts" is the powerful concept of photovoltaic powei systems Now a days photovoltaic systems are rapidly expanding and in creasing roles in electric power technologies, providing more secure power sources [25] The photovoltaic systems are, by nature, non linear power sources, that need accurate estimation of the maximum power generation for the operation as well as trading in deregulated powei market

In [12] estimation of photovoltaic power by using neural networks is developed, which uses environmental factors such as irradiation, temperature and wind ve locity for the prediction of photovoltaic power generation In [13] the effect of clouds on photovoltaic power generation is studied

The accurate prediction of the maximum power from the Photovoltaic power sys tems is inevitable, because of high uncertainty of solar radiation The power from the Photovoltaic systems depends on environmental factors, mainly the iiradia tion and the cell temperatuie It is very hard to measure the cell temperature, but related factors, like the temperature at the solar module and the wind ve locity can be used instead In this study these factors along with the irradiation will be used foi the forecast

For the prediction of maximum power generation, a three layer artificial neuial network is proposed The input signals are the irradiation, the temperature, and the wind velocity The output signal is the predicted maximum power from the photovoltaic system For the training data sets of the proposed neural network, two data sets, one for summer and other for winter are selected The neural

network has been trained by using one data set corresponding to one season, therefore, two types of the neural networks has been obtained from those data sets The accuracy of the prediction has been evaluated by using the actual data measured on the photovoltaic system The proposed model is also compared with conventional multiple regression analysis

This chapter is divided into five sections in which section 4 2 discusses the basics of the photovoltaic power generation and solar cell characteristics Section 4 3 deals with the forecasting methodologies mainly basics of the forecasting techniques, proposed neural network model and regression model In section 4 5 results and comparisons are discussed and in section 4 6 conclusion about the proposed model with its accuracy is mentioned

4 2   Fundamental Concepts of the Photovoltaic Modules

Before the description of the forecasting approach the fundamental concepts of photovoltaic power generation and solar cell operation will be discussed

4 2 1   Photovoltaic Power Generation

The generation of electricity from solar energy can be achieved through two major technology alternatives One uses the light from the sun to generate electricity directly (photovoltaic technologies), and the other uses the heat from the sun to increase the temperature of a working fluid, which-in turn can be used to generate electricity (solar thermal technologies) Each of these major alternatives can, in turn, be subdivided into variants of the major technology Photovoltaic technologies fall into crystalline, multi crystalline, thin film or concentrator variants while the solar thermal technologies fall into trough, power tower, dish engine and thermal electric variants

Photovoltaic power generation process is shown in figure 4 1 Generally speaking, photovoltaic solar cells use a semiconductor material that is exposed to sunlight The energy of the incident light displaces electrons from their normal atomic or
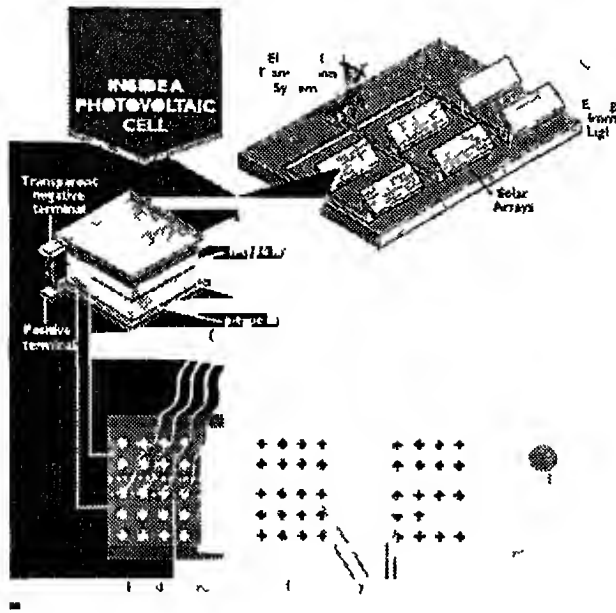
Figure 4 1   Photovoltaic power generation process

bits and an electrode grid structure on the surface of the semiconductor, collects these electrons and makes them available for use in an external circuit   This is very similar to the way that the chemical reaction and the electrodes in a dry battery cell make electrons available for external use   The terms crystalline, thin film and concentrator describe the manner in which the semi conducting material is processed and optimized as a photovoltaic cell   Crystalline cells are fabricated from ingot of the semiconductor material, usually silicon, that are cut into rel atively thin slices, processed to optimize the electron collection efficiency and laminated into a protective enclosure   Thin film cells are extremely thin layers of semi-conducting material that are evaporated onto a substrate, and concentrat ing cells use a plastic lens to concentrate sunlight from a large area onto a much smaller area of crystalline semi conducting material   All types have their merits and problems and are described in detail in [27]   Now after getting idea about the photovoltaic power generation, one have to understand about the solar cell characteristics which is described in next section
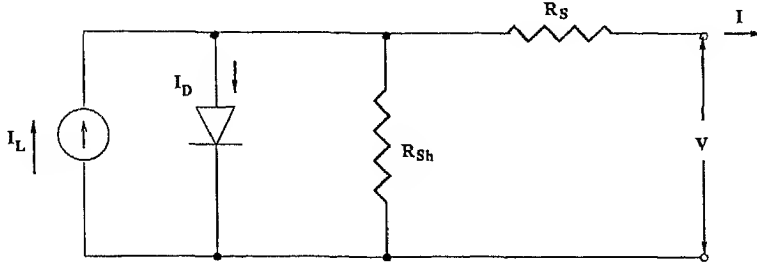
Figure 4 2   Equivalent circuit for a solar cell

4 2 2   Solar Cell Characteristics

A solar cell is simply a diode of large area forward bias with a photo voltage The photo voltage is created from the dissociation of electron hole pairs created by incident photons within the built in field of the junction or diode The electrical behavior of the cell is describable by the simple equivalent circuit shown in figure 4 2   Here the current-voltage characteristics for a solar cell are give by [14],

$$I \;=\; I_L - I_D - \frac{V + R_s I}{R_{sh}} \tag{4 1}$$

$$\text{where } I_D \;=\; I_o \left(exp \left(\frac{q(V + R_s I)}{nkT} \right) - 1 \right) \tag{4 2}$$

$$\;=\; I_L - I_o \left(exp \left(\frac{q(V + R_s I)}{nkT} \right) - 1 \right) - \frac{V + R_s I}{R_{sh}} \tag{4 3}$$

Where I_L is the photo generated current in Amperes

I_D is the ideal diode current in Amperes

I_o is the reverse saturation current of the diode in Amperes

q is the electronic charge in Coulomb

k is the Boltzmann constant in Joule per Kelvin

T is the junction temperature in Kelvin

R_s is the series resistance in ohms

R_{sh} is the shunt resistance in ohms

n is the ideality factor(the value of which depends on the recombination processes that are active in the cell)
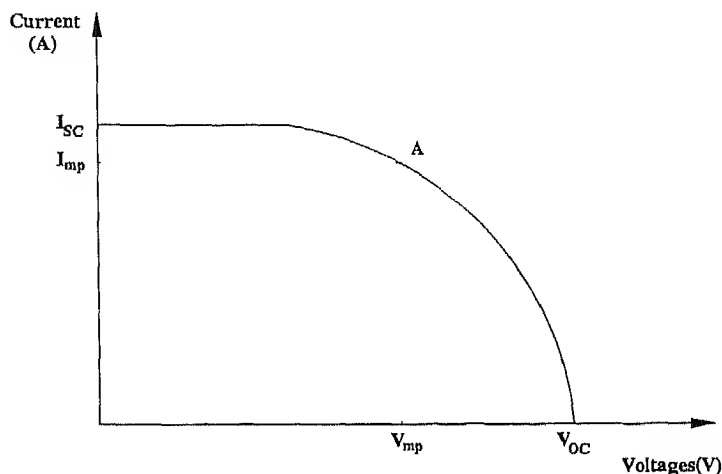
Figure 4 3   I   V Characteristics of a solar cell

An idealized current voltage characteristic of solar cell is shown in figure 4 3 One can note from the figure that the maximum output power of the solar cell is represented (A) the area of the maximum rectangle that can be drawn under the I V curve The point (A) usually referred to as the "knee" of the I-V curve The electrical characteristics of solar cells are based on their I V curves The I V curve is based on the cell being under standard conditions of sunlight and cell temperature, and assumes there is no shading in the cell Since photovoltaic cells are electrical semiconductors, partial shading may cause the cell to heat up Under this condition, the cells act as an inefficient conductor rather than an electrical generator Partial shading may ruin shaded cells and also affect the power output of the cell Figure 4 4 shows the I-V characteristics of a shaded and unshaded cell

The radiation and temperature are the main factor which effects the maximum power point of a solar cell The effect of radiation and temperature on maximum power point of a solar cell is shown in figure 4 5 As indicated in figure 4 5, the maximum power point increases, if the radiation increases or if the temperature decreases [14]

Some of the important definitions, which describes the cell properties found in
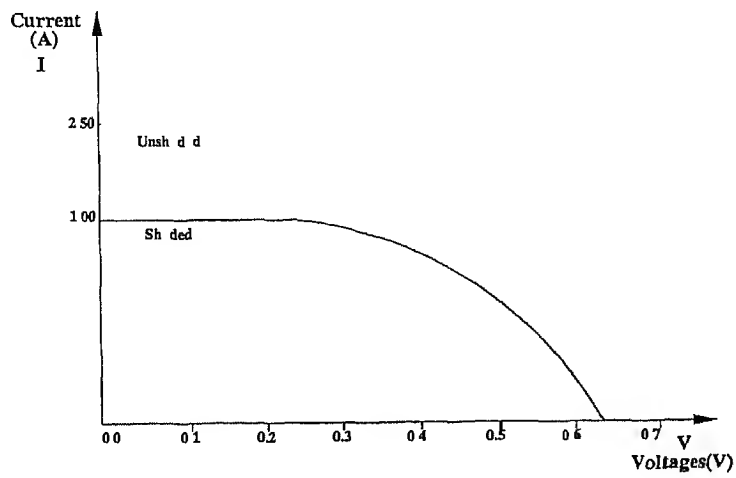
Figure 4 4   I   V Characteristics of a shaded and unshaded solar cell
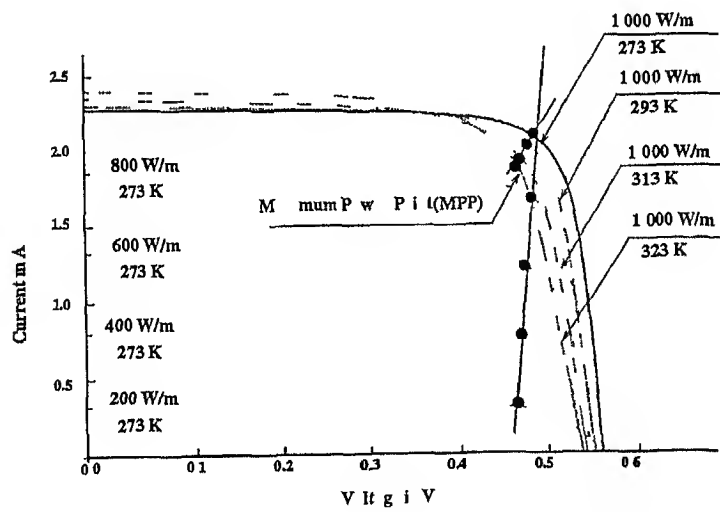
Figure 4 5   Effect of temperature and radiation on I   V Characteristics of a solar cell

literature [27] are presented here

Fill Factor(FF)   The ratio of the peak power to the product of the open circuit voltage (V_{oc}) and short circuit current (I_{sc}) is defined as the fill factor(FF)

$$FF = \frac{V_{mp}I_{mp}}{V_{oc}I_{sc}} \qquad (4\ 4)$$

where I_{mp} and V_{mp} are the current and voltage at maximum power point

Conversion Efficiency   The conversion efficiency is the ratio of the optimal electrical power (P_{opt}) delivered by the photovoltaic module to the solar insolation (E_e), received at a given cell temperature, T

$$\eta = \frac{P_{opt}}{AE_e} \qquad (4\ 5)$$

Where the optimal power $P_{opt} = V_{oc}I_{sc}FF$ is in Watts, E_e is in Watts per square meter and the cell area, A, is in square meters

4 3   Forecasting Methodologies

According to the basics of photovoltaic power generation and solar cell characteristics, one can say that photovoltaic power generation mainly depends on the irradiation and the cell temperature It is very difficult to measure cell temperature, so instead of that, weather parameters like temperature and wind velocity are utilized In this study irradiation, temperature and wind velocity are utilized for forecasting photovoltaic power

Here two forecasting methodologies are presented which were used by the author The first one is based on artificial neural networks The second one is based on multiple regression and statistical analysis

4 3 1   Neural Network Based Prediction Model

The artificial neural network used in this study, is a multi layered feed-forward network using a back propagation training algorithm The detailed explanation of this kind of network is explained in chapter 6 Here the configuration of proposed model for estimating the photovoltaic power and its training process is explained
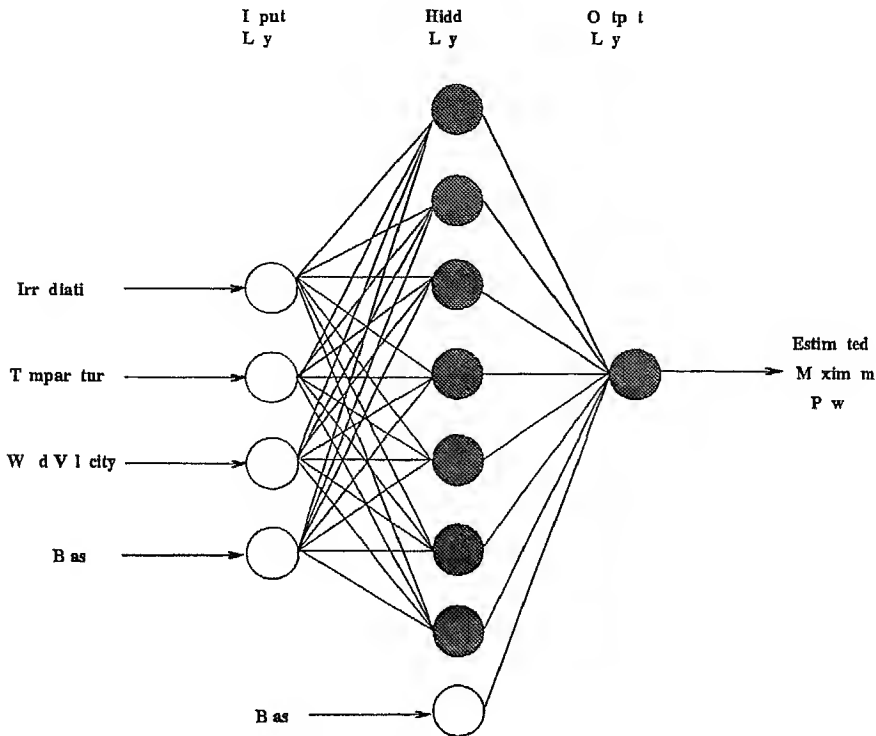
Figure 4 6   Configuration of the neural network

4 3 1 1   Configuration of Artificial Neural Network

The configuration of proposed neural network is shown in figure 4 6   The network has three layers, namely input layer, hidden layer and output layer   The input layer has four nodes for irradiation, temperature , wind velocity and a bias signal of 1   The hidden layer has eight nodes including one bias node   The output layer has only one node, and the output signal is the predicted maximum power   The nodes are classified into two groups as white nodes and grey nodes   In the white nodes no activation function is used, which means that the input goes directly to the output   In the grey nodes unipolar continuous sigmoidal activation function (see chapter 6) is used   The bias nodes are used here to improve the learning speed in the training process [12]

4 3 1 2   Initialization of The Training

In all types of neural networks the training should be started by using proper weight matrices for output and hidden layers   Here the weight matrices are chosen randomly in between [0 5 0 5] except 0, for both output and hidden layers   Before the training process the input vector should be normalized   Here the input vector is normalized in between 0 1 to 0 9 by using the below equation where x_{ori} and x_{norm} are the original and normalized input values

$$x_{norm} = \frac{0\ 8}{max(X) - min(X)}(x_{ori} - min(X)) + 0\ 1 \qquad (4\ 6)$$

where X is consider the input vector   This normalization is to give the 10 percent variation of the minimum and the maximum range of the input vector   The completion of the training process is done by checking the mean square error (MSE)   This can computed by the equation below

$$MSE = \frac{1}{N_P} \sum_{N_P} \sum_{l} (d_l - y_l)^2 \qquad (4\ 7)$$

where N_P is the number of training patterns, l is the number of output neurons, in this case it is one d_l and y_l are the desired and actual outputs   The training process completes if the MSE will be less than or equal to the specified value ϵ Here ϵ is chosen as 0 001   If this is not met then one should give the maximum number of iterations ($ITER$), that the neural network should be trained   The learning rate (η) and the momentum (β) are chosen 0 5 and 0 3   During the training learning rate is changed as explained in chapter 6, section 6 5

4 3 2   Regression Based Prediction Model

The regression model used in this study, is simple linear regression model with multiple inputs   The detailed analysis of linear regression with multi variables is explained in chapter 5   The linear regression equation for photovoltaic power generation can be written as shown below

$$Power = a + b_1\,Rad + b_2\,Temp + b_3\,Ws \qquad (4\ 8)$$

Where Rad is the irradiation $Temp$ is the temperature and Ws is the wind speed

By using regression analysis, the above regression coefficients a, b_1, b_2 and b_3 are to be found

4 4   Data Analysis

The data that is used for performing neural network and multiple regression forecasting models was taken form an experimental solar cell configuration which belongs to the Institut fuer Elektrische Energieversorgung, Technical University, Darmstadt, Germany   In that configuration there are twelve modules and each module is having 36 solar cells

The solar cells are of *mono crystalline* type and they are connected serially in this configuration   These solar cells are from Siemens and the specification type is $M55S$ Some main characteristics like peak power, short circuit current and open circuit voltage are presented below for one photovoltaic module

Maximum Power (P_{max})	53 0 (W)
Short Circuit Current (I_{SC})	3 27 (A)
Open Circuit Voltage (V_{OC})	21 8 (V)

Table 4 1   Specifications of M 55 S photovoltaic module

In this configuration measurements were taken for each day with one second interval   The measurements are carried for radiation, temperature and wind velocities, were taken 24 hours for each day with one second interval   The data used in this model are from January and June 1994   January data is for winter period power forecasting and June data is for summer period power forecasting In the summer periods, the data period is considered from 3 a m to 9 p m   During the winter season the data period is considered from 8 a m to 6 p m   The noise in the data was filtered and modified before it was used in the models   Through the modification of the data, the values were reduced to one value for each minute This was done by taking the average for each minute interval data

4 4 1   Selection of Training Data Sets

For the training of the proposed neural network training the data sets were divided into two groups   One is summer data set and other is winter data set   As explained in previous section, summer and winter data are having significant variations, so it is better to get two neural network structures separately for summer and winter data sets   To get proper neural network structure in each season, three consequent days are selected in each season   The training for each season is carried in two phases   In the first phase the middle day is used to train the net and other two days are used to test the network   In second phase the first two days are used to train the net and third day is used to test the net   The data sets that are used in summer and winter power forecasting are given below

Summer data sets   June 20th ,21st and 22nd of 1994, are selected to get proper neural net for summer period   These figures are shown in Appendix A 5, A 6 and A 7   In these figures power, radiation, temperature and wind velocity are plotted with respect to time   In phase I 21st day data is selected to train the net and 20th and 22nd days are used for testing   In phase II 20th and 21st days are used to train the net and 22nd day data used for testing

Winter data sets   January 9th, 10th and 11th of 1994, are selected for winter power forecasting   These figures are also shown in Appendix A 8, A 9 and A 10   In phase I 10th day data is chosen for training purpose, and 9th and 11th data's are used for testing   In phase II combinedly 9th and 10th data is used to train the net and 11th day data set is used for testing the accuracy of the neural net

The training and testing results of both the neural networks for summer and winter are discussed in next section
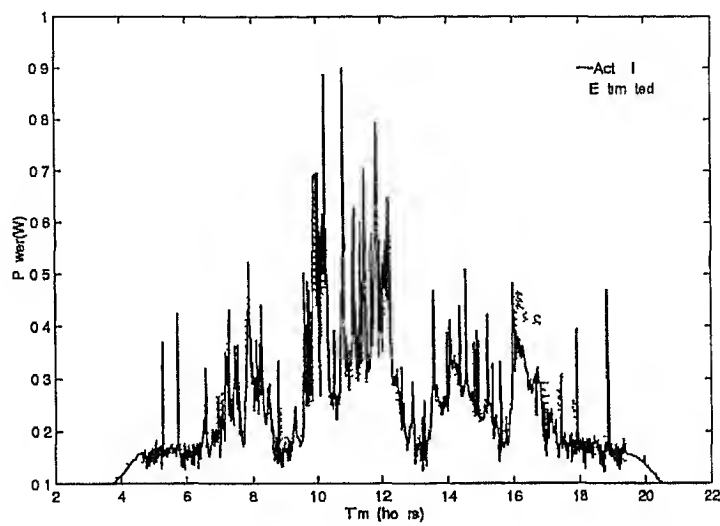
Figure 4 7   Phase I 22nd june data neural network results

4 5   Results and Comparisons

4 5 1   Forecasting Results by Using Neural Networks

Training of the proposed neural network model is carried for summer and winter seasons in both phases separately For the summer period for phase I the neural network results are shown in figure 4 7 For the better view of the estimation accuracy two specific hours data was chosen for testing the proposed model, they are 14th hour of 20th june data and 9th hour of 22nd june data The results of these two hours are presented in Appendix A 13 and A 14 Similarly for the second training phase results are shown in figure 4 8 and for better view, 9th hour of that day is chosen for testing and is shown in Appendix A 15 The results corresponding to the 20th june were also presented in Appendix A 12 The accuracy test of the proposed model is done by calculating the absolute percentage error that is

$$APE = \frac{|AC - ES|}{AC} \times 100 \qquad (4\ 9)$$
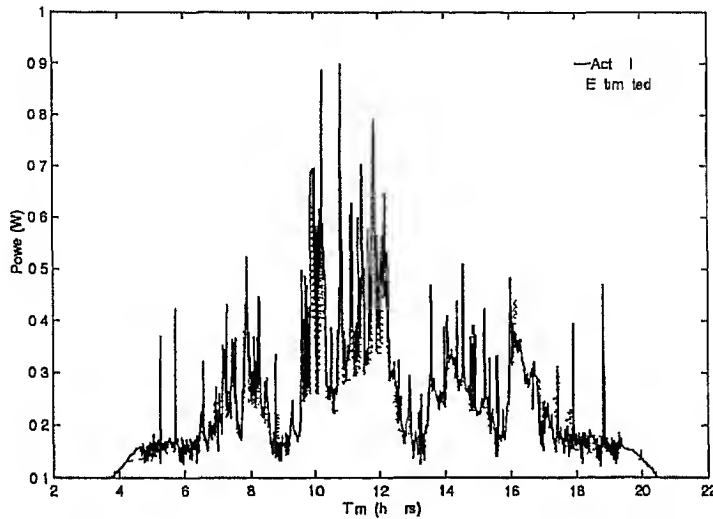
Figure 4 8   Phase II 22nd june data neural network results

Where AC is the actual measured value and ES is the estimated value of the power

In the similar manner it is carried for the winter data also   The results for the winter season is shown in figure 4 9 for the phase I neural network   For phase II the results are shown in figure 4 10

4 5 2   Forecasting Results by Using Multiple Regression

In Regression analysis the data is also divided into two seasons and for each season the analysis process is carried in two phases   For summer period, in phase I, 21st june data is taken for regression analysis and tested with other two days   In phase II total two days that is 20th and 21st are chosen for regression analysis and is tested for the third day   The statistical analysis of the data is carried for 20th june, 21st june and for both which are shown in Table 4 2   The value of R^2, which is called multiple regression coefficient, is nearer to 1, for each day and the total of the both days, is very encouraging (see chapter 5)

It means that the analysis is very good and it can estimate the other day power more accurately   The forecasting results are shown in figure 4 11 for the phase I
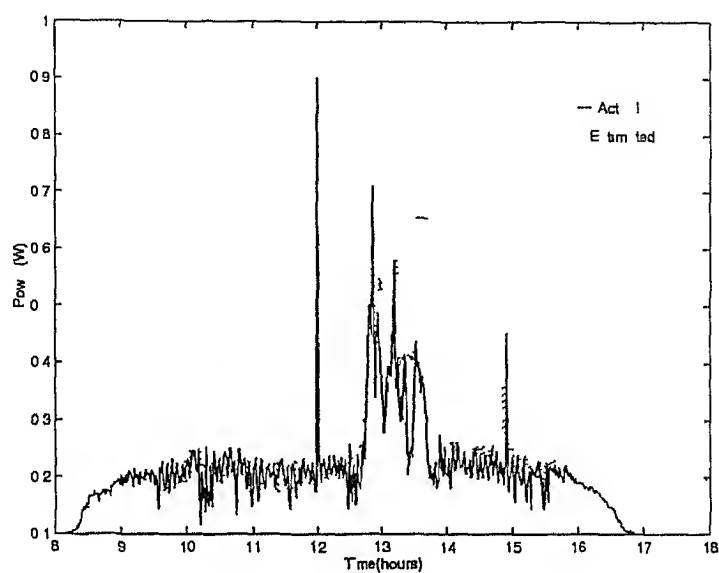
Figure 4 9   Phase I 11th january data neural network results
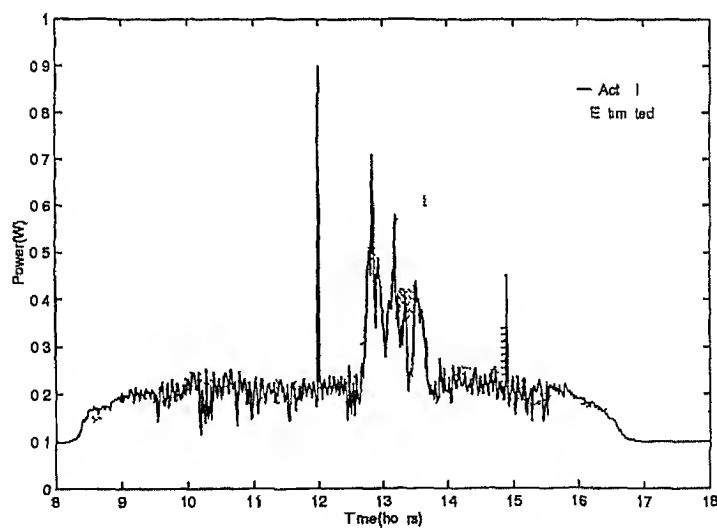
Figure 4 10   Phase II 11th january data neural network results
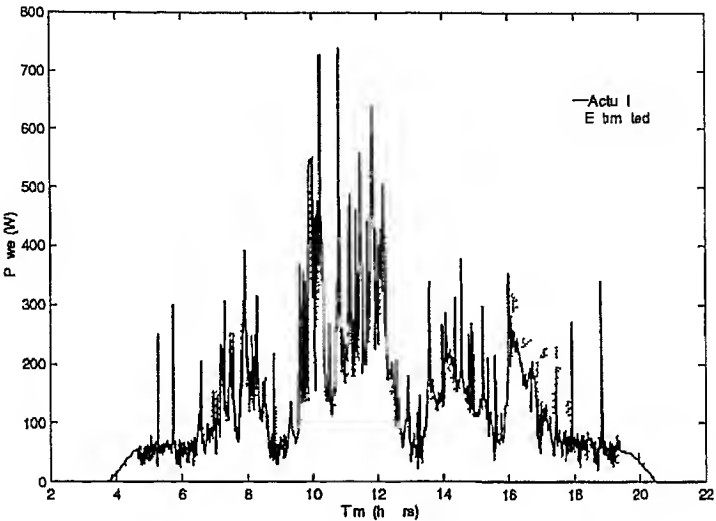
Figure 4 11   Phase I 22nd june data multiple regression results
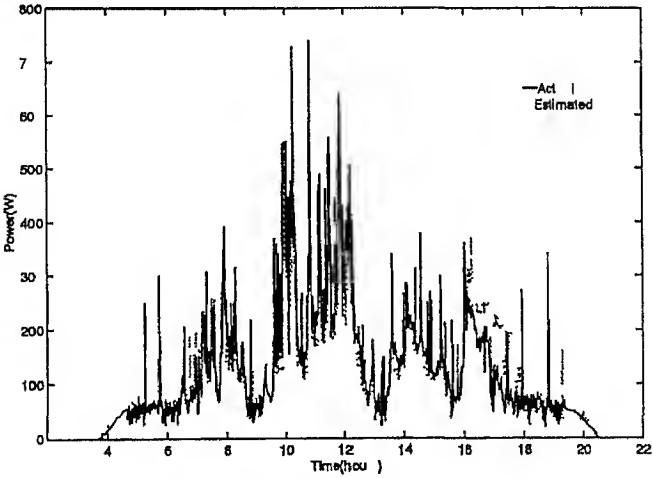
Figure 4 12   Phase II 22nd june data multiple regression results

Data File	Intercept (a)	b_R	b_T	b_W	R^2	R
20th June	0 0604	0 8788	0 11841	0 05608	0 946528	0 97289
21st June	0 05804	0 82299	0 04062	7 198E-4	0 94329	0 971235
Total	0 05677	0 8401	0 08203	0 05774	0 94404	0 971619

Table 4 2   Statistical results for summer data

and for phase II it is shown in figure 4 12   Again for the better view of the results hourly data is also chosen and is shown in Appendix figure A 18 and figure A 19 for phase I and figure A 20 for phase II

Variable	b	S_b	F_b	Correlation with Power
Radiation	0 82299	0 007645	11589 95	0 97079
Temperature	0 04062	0 011439	12 60846	0 57665
Wind Speed	7 198E-04	0 02543	8 013E-4	0 033435

$R^2_{P(T\,Ra)} = 8\,642\text{E-}004$	$R_{P(RaT)} = 0\,971235$	$R^2_{P(W\,RaT)} = 0\,000000$
$F_{R^2} = 5966\,74$		

Table 4 3   21st June data multiple regression analysis

The statistical results for 21st and total of 20th plus 21st are presented in Table A 1 and Table 4 4   The significance of R^2 and also the regression coefficients are tested by the F test ($F_{R^2}$and F_b) and the results are presented in those tables, which are good   Test of significance of R^2 and regression coefficients is discussed in chapter 5, section 5 8   The results of significance test are comparable with the null hypothesis value (see chapter 5)   The standard error's of the regression coefficients(S_b) and semi partial correlation coefficients($R^2_{P(T\,Ra)}$ and$R^2_{P(W\,TRa)}$) are also presented with $R_{P(TRa)}$ The standard error of regression coefficients is less which means the regression equation is better fit for the historical data   The semi partial correlation coefficients tells about the effect of each independent variables, that is radiation, temperature and wind speed, when they are presented in the regression analysis sequentially

Variable	b	S_b	F_b	Correlation with Power
Radiation	0 8401	0 00597	19790 8	0 97006
Temperature	0 08203	0 00787	108 548	0 62745
Wind Speed	0 05774	0 015198	14 4322	0 360015

$R^2_{P(T\,Ra)} = 0\ 0026$	$R_{P\,(RaT)} = 0\ 971421$	$R^2_{P(W\,RaT)} = 3\ 8472\text{E-}004$
$F_{R^2} = 11787\ 448$		

Table 4 4   20th + 21st June data multiple regression analysis
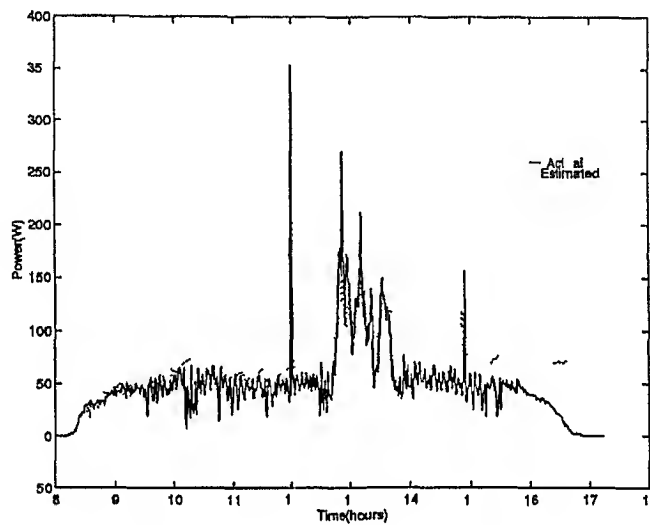
Figure 4 13   Phase I 11th january data multiple regression results

Similar procedure is applied for the winter data also and forecasting results are shown in figuie 4 13 corresponding to phase I and figure 4 14 corresponding to phase II The results for the 9th january data were shown in Appendix figure A 16   From the statistical analysis the values of R^2 for all the days are very fai from 1   It means that the data is not good to fit the regression equation Although significant test results shows very fai from the null hypothesis value (see chapter 5) , the standard error of the iegression equation not worth mentioned

Statistical analysis of the winter data that is 9th, 10th and total (9th+10th) is ulso cairied which are shown in Table 4 5   Individual statistical results for 10th and total (9th+10th) are presented in Table 4 6 and Table 4 7   Statistical results
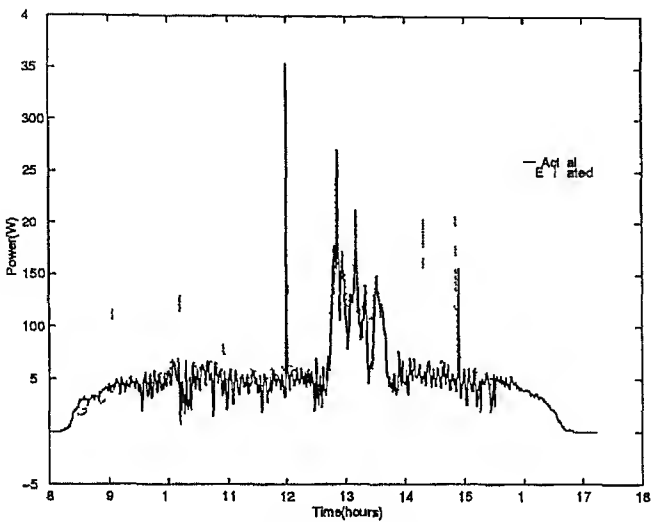
Figure 4 14   Phase II 11th january data multiple regression results

for the 9th january is shown in Appendix A

Data File	Intercept (a)	b_R	b_T	b_W	R^2	R
9th January	0 00856	0 91547	0 12543	0 00863	0 6695	0 81822
10th January	0 02861	0 25366	0 30732	0 03667	0 42369	0 65091
Total	0 01258	0 61349	0 24722	-0 02646	0 57422	0 75778

Table 4 5   Statistical Results for winter data

Variable	b	S_b	F_b	Correlation with Power
Radiation	0 25365	0 03212	62 37185	0 5118
Temperature	0 30732	0 2377	167 152	0 60166
Wind Speed	0 03667	0 02099	3 0508	0 06176

$R^2_{P(T\,Ra)} = 0\ 1588$	$R_{P(RaT)} = 0\ 64864$	$R^2_{P(W\,RaT)} = 0\ 0029$
$F_{R^2} = 146\ 0576$		

Table 4 6   10th January data multiple regression analysis

Variable	b	S_b	F_b	Correlation with Power
Radiation	0 61349	0 02795	481 7477	0 72705
Temperature	0 24722	0 02235	122 3128	0 63455
Wind Speed	0 02646	0 014185	3 47998	0 02444

$R^2_{P(T\,Ra)} = 0\,0444$	$R_{P(RaT)} = 0\,75695$	$R^2_{P(W\,RaT)} = 0\,0013$
$F_{R^2} = 537\,6658$		

<div align="center">Table 4 7   9th + 10th January data multiple regression analysis</div>

4 5 3   Comparisons

The comparison study between neural network model and regression model is explained through the the following tables

For summer period photovoltaic power forecasting, both the models gave better results in both phases   The phase I and phase II comparison study results are shown in Table 4 8 and Table 4 9   The absolute percentage error for both models are below 10% which shows us the results are better and accuracy of estimation is good

Actual Power(W)	Estimated Power by regression	Error (AC ES)/AC	Estimated Power by neural net	Error (AC - ES)/AC
285 933	273 1053	0 04486271	297 03376	0 03882295
248 11957	242 82426	0 02134174	263 19864	0 06077338
313 97864	304 26276	0 0309444	329 74286	0 05020795
367 62003	342 24298	0 06903064	372 61959	0 01359981
421 95712	392 04013	0 07090055	421 06262	0 00211988
281 7306	275 64465	0 02160198	291 26263	0 03383386
521 8731	516 5613	0 01017839	534 28768	0 0237885
339 97058	327 42935	0 03688916	354 81731	0 04367064
33 43773	30 891485	0 07614882	32 954305	0 01445748
33 177418	31 382862	0 05408967	30 922437	0 06796735
31 31728	31 255587	0 00196994	30 737674	0 01850755
31 113144	31 867685	0 02425153	29 954991	0 03722391

<div align="center">Table 4 8   Summer forecasting results in phase I</div>

Actual Power(W)	Estimated Power by regression	Error (AC ES)/AC	Estimated Power by neural net	Error (AC ES)/AC
285 933	268 78052	0 05998781	272 63153	0 04651954
248 11957	241 03696	0 02854515	240 38665	0 03116611
313 97864	298 01813	0 05083311	303 5089	0 03334538
367 62003	339 5232	0 076429	352 07792	0 04227764
421 95712	391 26404	0 07273982	401 64712	0 04813286
281 7306	276 10715	0 01996036	268 8079	0 04586901
521 8731	508 85815	0 02493892	500 78032	0 04041745
339 97058	323 47888	0 0485092	332 11618	0 02310317
33 43773	32 367237	0 03201449	32 195081	0 03716307
33 177418	33 57368	0 01194376	32 115856	0 03199652
31 31728	31 322279	0 00015963	30 441757	0 02795654
31 113144	33 97892	0 09210821	31 88181	0 02470552

Table 4 9   Summer forecasting results in phase II

For the winter period photovoltaic power forecasting, it is found from the results that the neural network model is better than the regression based model   But comparing with the summer period forecasting results, the winter period fore casting results are found bad   This is because of high noisy in the winter period data   The results corresponding to winter data is shown in Table 4 10 and Table 4 11

Actual Power(W)	Estimated Power by regression	Error (AC − ES)/AC	Estimated Power by neural net	Error (AC - ES)/AC
22 45075	38 281857	0 7051482	20 607161	0 08211705
45 07448	63 420948	0 40702567	41 997104	0 06827314
39 700428	69 09451	0 74039716	42 920444	0 08110784
56 287014	74 07306	0 31598842	52 437921	0 06838333
49 0921	64 68476	0 31762064	45 463991	0 07390412
35 268646	55 965977	0 5868479	32 934294	0 06618773
46 01911	71 24653	0 54819435	47 656461	0 03557981
53 43403	73 73588	0 37994233	51 044439	0 0447204
102 58027	96 9917	0 05447997	94 89121	0 07495652
123 82854	105 58305	0 1473448	113 72289	0 08161005
178 26665	117 01831	0 3435771	165 59396	0 07108841
171 75298	121 07685	0 29505238	155 61626	0 0939531
69 930626	73 8231	0 0556619	72 018864	0 02986157
58 056194	76 871765	0 3240924	61 180204	0 05381011
67 82848	78 269226	0 15392858	69 057034	0 01811265

Table 4 10   Winter forecasting results in phase I

Actual Power(W)	Estimated Power by regression	Error (AC − ES)/AC	Estimated Power by neural net	Error (AC − ES)/AC
22 45075	25 763172	-0 1475417	23 754934	0 05809088
45 07448	60 269527	0 33710983	45 785301	0 01576992
39 700428	61 46312	-0 5481727	42 10977	0 06068805
56 287014	73 18864	-0 3002757	53 665741	0 04656976
49 0921	62 826622	0 27977055	50 609778	0 03091491
35 268646	46 833687	0 3279128	35 61555	0 00983603
46 01911	65 174706	0 416253	46 731944	-0 01548996
53 43403	70 86716	0 32625517	51 055351	0 04451618
102 58027	117 29827	0 1434779	103 03717	-0 00445408
123 82854	132 405	0 06926078	126 02267	0 01771912
178 26665	151 21524	0 15174687	181 4929	-0 01809787
171 75298	162 0112	0 05671969	180 92085	0 05337823
69 930626	74 128975	0 06003591	65 69936	0 06050662
58 056194	73 20502	0 2609338	57 035314	0 01758434
67 82848	76 70768	0 1309066	63 767411	0 05987263

Table 4 11   Winter forecasting results in phase II

4 6   Conclusions

In this chapter photovoltaic power forecasting models weie discussed Two mod els weie constiucted for photovoltaic power forecasting, a neural network based model and a multiple regression based model In neural network model standard multilayer feed foiward back propagation learning algorithm was used   Photo voltaic power forecasting is carried for two seasons, summer and winter Results weie presented for both the models coiresponding to summer and winter From the results it is concluded that the summer period power forecasting gave better results than the winter period power forecasting It is found due to the high noisy in wintei data comparing with summer data For the summer period both the models gave good results and for the winter period neural network is better than the iegiession model The used data included recognizable noise and it would be interesting to test the models with more accurate data

5 Multiple Regression

5 1   Introduction

The problem of finding the most suitable form of equation to predict one variable from the values of one or more other variables is called the problem of regression [7] In order to estimate the regression curve of y on x we must first specify the *functional form* of the curve Some examples are

$$Y = a + bX \tag{5 1}$$

$$Y = a + b_1 X + b_2 X^2 \tag{5 2}$$

$$Y = ae^X \tag{5 3}$$

Here only linear regression is considered with one Y variable and several X variables   The analysis of linear regression model for multiple X variables is known as Multiple Regression Analysis   Multiple Regression equation can be written as shown in the below equation

$$Y' = a + b_1 X_1 + b_2 X_2 +, \quad , + b_k X_k \tag{5 4}$$

One h ive to find the values of $a, b_1, b_2, \quad , b_k$ that will result in the highest possible positive correlation between the observed values Y and Y' This can be done by using historical data

5 2   Multiple Regression Analysis

With k different X variables the multiple regression equation is

$$Y' = a + b_1 X_1 + b_2 X_2 +, \quad , + b_k X_k \tag{5 5}$$

The total sum of squared deviations of the Y values from the mean of the Y values $(SS_{tot} = \sum(Y - \overline{Y})^2)$ can be partitioned into two orthogonal components the sum of squares for the linear regression $(SS_{reg} = \sum(Y' - \overline{Y})^2)$ and the residual sum of squares $(SS_{res} = \sum(Y - Y')^2)$ [8]

$$SS_{tot} = SS_{reg} + SS_{res} \tag{5 6}$$

$$\sum(Y - \overline{Y})^2 = \sum(Y' - \overline{Y})^2 + \sum(Y - Y')^2 \tag{5 7}$$

where $\overline{Y}$ is the mean of Y values   To find the values of $a, b_1, b_2, \quad , b_k$ the residual sum of squares should me minimized   Then the required value of a will be

$$a = \overline{Y} - b_1 \overline{X}_1 - \quad - b_k \overline{X}_k \tag{5 8}$$

Where $\overline{X}$ is the mean of X values   Substituting this value of a in 5 1, then

$$Y' = \overline{Y} + b_1(X_1 - \overline{X}_1) + b_2(X_2 - \overline{X}_2) + \quad + b_k(X_k - \overline{X}_k) \tag{5 9}$$

Then, if $\sum(Y - \overline{Y})^2$ is to be minimized, the values of b must satisfy the following equations

$$b_1 \sum(x_1^2) + b_2 \sum(x_1 x_2) + \quad + b_k \sum(x_1 x_k) = \sum(x_1 y)$$
$$b_1 \sum(x_2 x_1) + b_2 \sum(x_2^2) + \quad + b_k \sum(x_2 x_k) = \sum(x_2 y)$$

$$\tag{5 10}$$

$$b_1 \sum(x_k x_1) + b_2 \sum(x_k x_2) + \quad + b_k \sum(x_k^2) = \sum(x_k y)$$

wheie

$$x_i = X_i - \overline{X} \tag{5 11}$$

$$y_i = Y_i - \overline{Y} \tag{5 12}$$

The values of b can be obtained by solving the above equations After getting the values of b they can be substituted in 5 8 to find a The meaning of a regression coefficient in a multiple regression equation is If all variables except X_i are held constant, then b_i is the amount by which Y_i increases with unit increase in X_i [8]

5 3   The Correlation Coefficient

The correlation coefficient between two variables Y and X can be defined as

$$r = \frac{\sum xy}{\sqrt{\sum x^2 \sum y^2}} \tag{5 13}$$

The significance and the importance of the correlation coefficient can be ex plained by analyzing the equation 5 2

In case of two variables Y and X, if there is a perfect linear relationship be tween Y and X so that all of the plotted points fall precisely on a straight line, then SS_{res} will be equal to zero and SS_{reg} will be equal to SS_{tot} If there is no tendency for the Y values to be linearly related to the X values, then the regression coefficient B and SS_{reg} will be equal to zero and SS_{res} will be equal to SS_{tot} An important index of the degree to which the Y and X values are linearly related is obtained by dividing both sides of the above equation by SS_{tot} [8]

$$\text{then} \quad 1 = \frac{SS_{reg}}{SS_{tot}} + \frac{SS_{res}}{SS_{tot}} \tag{5 14}$$

$$\text{putting} \quad \frac{SS_{reg}}{SS_{tot}} = R^2_{YY'} \tag{5 15}$$

$$\text{then} \quad 1 = R^2_{YY'} + (1 - R^2_{YY'}) \tag{5 16}$$

where $R_{YY'}$ is the correlation coefficient or in case of multiple regression it is called Multiple regression coefficient Obviously R^2_{YY} cannot be greater than 1
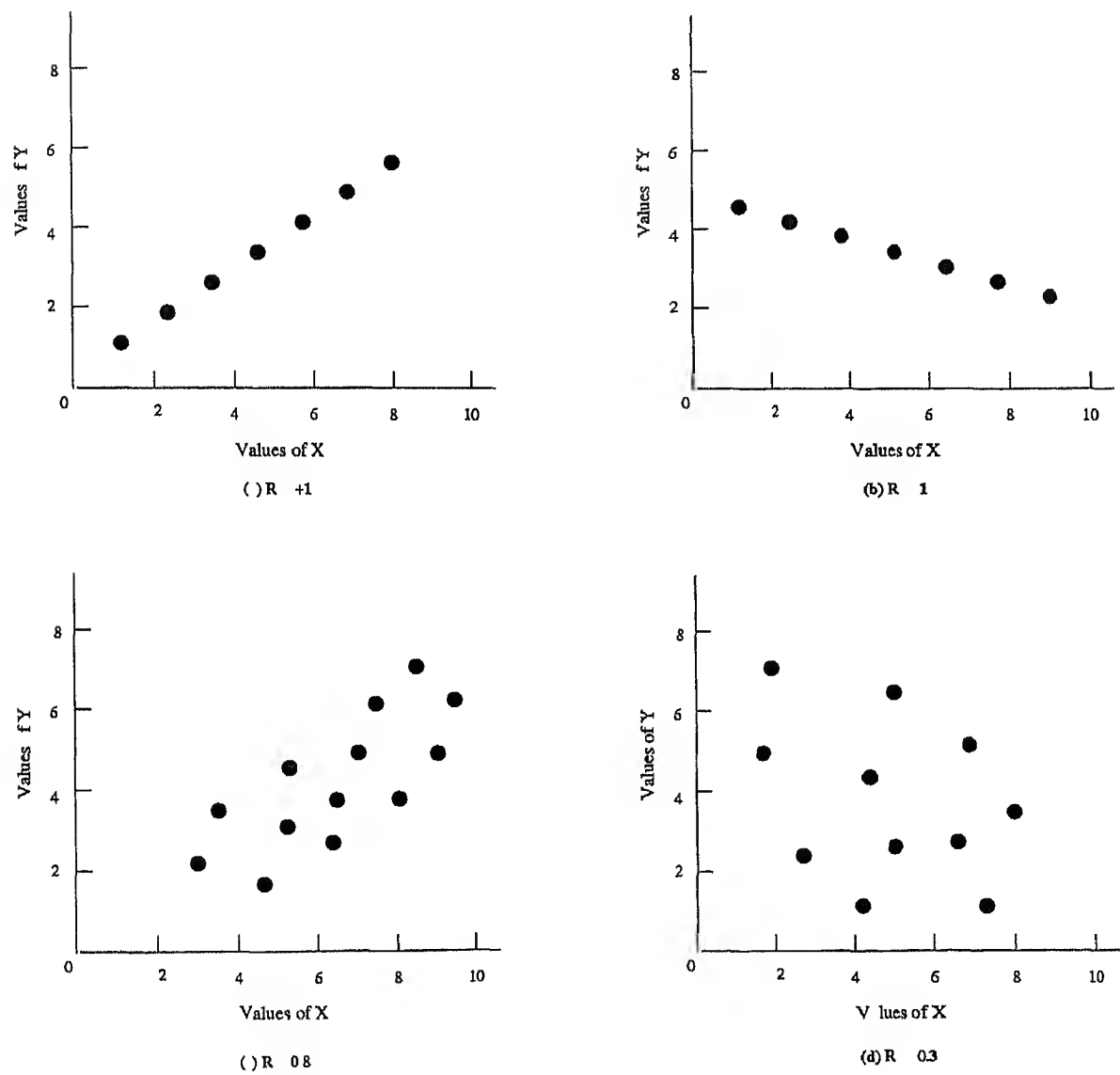
Figure 5 1   The significance of Correlation Coefficient

and consequently, R_{YY} can range only within the limits 1 to 1   The significance of R_{YY} will be explained in case of one input (X) and one output (Y) below [8]

In case of two variables that is X and Y the significance of r is explained by using four graphs shown in figures a,b,c,d   Figure 5 1a and figure 5 1b are the plots where $r = 1\ 00$ and $r = \ 1\ 00$, respectvly   In figure 5 1c, r value approximately 0 8 and in figure 5 1d it is close to 0 3   The value of r^2 provides an index of the degree to which a set of plotted points cluster about the regression line   The closer the points fall along the regression line, the larger the value of r^2 and the greater the proportion of the total sum of squares accounted for by the linear regression of Y on X When the value of r^2 is small as shown in Figure 5 1d, the plotted points will show considerable scatter about the regression line, and the proportion of the total sum of squares accounted for by the linear regression of Y on X will be small

5 4   Regression Analysis with Standardized Variables

If all the variables in multiple regression problem that is k variables of X and Y are in standardized form, then the value of a in the regression equation will be equal to zero [8]   In this instance all of the regression coefficients will be in standardized form, and the regression equation will take the form

$$z_{Y'} = b_1 z_1 + b_2 z_2 + \quad + b_k z_k \tag{5 17}$$

The definition of *standardized variable* is one that has a mean value equal to zero and a variance and standard deviation equal to 1   Any variable Y can be transformed into a standardized variable by means of the following linear transformation

$$z = \frac{Y - \overline{Y}}{s_Y} \tag{5 18}$$

where s_Y is the standard deviation of Y values

If the residual sum of squares is to be minimized, then the regression coefficients

must satisfy the following equations

$$b_1 r_{11} + b_2 r_{12} + \quad + b_k r_{1k} = r_{Y1}$$
$$b_1 r_{21} + b_2 r_{22} + \quad + b_k r_{2k} = r_{Y2}$$

$$(5\ 19)$$

$$b_k r_{k1} + b_2 r_{k2} + \quad + b_k r_{kk} = r_{Yk}$$

where r_{ij} is the inter correlation coefficient between x_i and x_j, and r_{Yi} is the correlation coefficient between Y and x_i for $i = 1,2,3, \quad$ k   The $k \times k$ inter correlations of the X variables can be represented by a symmetric matrix which is designate by R_{ij} The k correlations of the X variables with the Y variable can be represented by a $k \times 1$ column vector that is designate by R_{Yi} Similarly, the k values of the regression coefficients can be represe noted by a $k \times 1$ column vector that is designate by B_i To find B, find the inverse of R_{ij} and multiply with R_{Yi}

$$R_{ij}^{-1} R_{Yi} = B_i \qquad (5\ 20)$$

One can obtain the original multiple regression coefficients that is B by using the below formula

$$b_i = \frac{s_1}{s_X} b_i \qquad (5\ 21)$$

From these values of b, one can obtain a by using 5 8

5 5   Mean Squares

When a sum of squares is divided by its degrees of freedom (d f), the resulting value is called a mean square The number of degrees of freedom can be thought of as the number of independent comparisons available [7] From the literature [8] the degrees of freedom of each sum of squares *that is* SS_{tot}, SS_{reg}, and SS_{res}, is $n-1$, k and $n-k-1$ Where n is the number of samples, k is the number of X variables Then the mean squares of each sum of squares is

$$MS_{tot} = \frac{SS_{tot}}{n-1} \tag{5 22}$$

$$MS_{reg} = \frac{SS_{reg}}{k} \tag{5 23}$$

$$MS_{es} = \frac{SS_{res}}{n-k-1} \tag{5 24}$$

5 6   Standard Error of the Regression Coefficients

The standard error of a regression coefficient b_i when several X variables are involved will be given by

$$S_b = \sqrt{\frac{MS_{res}}{\sum x_i^2(1-R_i^2)}} \tag{5 25}$$

where R_i^2 is the squared multiple correlation of X_i with the remaining X variables When the inverse of the correlation matrix between the X variables is available, the value of R_i^2 is easily obtained It will be given by

$$R_i^2 = 1 - \frac{1}{r^{ii}} \tag{5 26}$$

where r^{ii} is the diagonal entry in the inverse of the correlation matrix

5 7   The Semi Partial Correlation Coefficient

A semi partial correlation coefficient is the correlation coefficient between two variables X_i and X_j after the variance that X_k has in common with X_i and X_j is removed from only one of the two variables In case of three variables, the

semi partial correlation coefficient between X_1 and X_2 with the variance of X_3 removed from X_2 will be given by

$$r_{1(2\ 3)} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{1 - r_{23}^2}} \qquad (5\ 27)$$

where the notation $r_{1(2\ 3)}$ indicates that it is the correlation coefficient between X_1 and X_2 after the variance that X_3 has in common with X_2 has been eliminated from X_2 Applying this for one Y and two X variables, semi partial correlation coefficient can be written as

$$r_{Y(1\ 2)} = \frac{r_{Y1} - r_{Y2}r_{12}}{\sqrt{1 - r_{12}^2}} \qquad (5\ 28)$$

The relation between semi partial correlation and correlation coefficient is

$$r_{Y(1\ 2)}^2 = R_{Y\ 12}^2 - r_{Y2}^2 \qquad (5\ 29)$$

Generalizing this for any number of X variables, then correlation coefficient can be written as

$$R_{Y\ 123\ \ k}^2 = r_{Y1}^2 + r_{Y(2\ 1)}^2 + \quad + r_{Y(k\ 123\ \ k-1)}^2 \qquad (5\ 30)$$

The theoretical behavior of above equation can be explained for three X variables as follows   The proportion of the total sum of squares accounted for by X_1, given that it is entered first in the regression equation, will be r_{Y1}^2 After X_1 is entered in the regression equation, the increment in the proportion of the total sum of squares accounted for by X_2, given that it is entered second in the regression equation, will be $r_{Y(2\ 1)}^2$ Now that both X_1 and X_2 are in the regression equation, the increment due to X_3 will be given by $r_{Y(3\ 12)}^2$ [8]

5 8   Test of Significance

One of the most important thing in the multiple regression analysis is "test of significance" of multiple regression equation coefficients and correlation coefficient   Significance tests for these values are carried by using statistical tests   Commonly known tests in the statistics are F test, t test and χ^2 test [7]

General procedure for making some test is explained here with an example [7] Assume there are n samples in the sample space which are taken from a normal distribution with unknown mean μ. The sample mean is denoted by $\bar{x}$. If one is interested in a particular value μ_o and asks some question 'does $\bar{x}$ differ significantly from μ_o?', then test of significance can be carried by calculating the observed "t test' value. It can be found by using below equation

$$t_o = \frac{\bar{x} - \mu_o}{s/\sqrt{n}} \tag{5 31}$$

Where s stands for the standard deviation of the samples. Now the percentage point $t_{\alpha/2\ v}$ is chosen so that there is a probability α of getting a larger observation from a t distribution with v degrees of freedom. If $|t_o|$ is larger than $t_{\alpha\ v}$, then the result is significant at the 100α percent significance level. The values of $t_{\alpha\ v}$ is available in most of the statistical books [7]

Similar test procedure is taken for the *correlation coefficient* and *regression co efficients*. As the above procedure is explained for the t test, same procedure is also carried for the F test also. But the use of F test is it compares more than one sample space. In this case the chosen percentage point can be denoted by $F_{\alpha/2\ v1\ v2}$, where $v1$ and $v2$ are the degrees of freedom for the two samples

5 8 1   Test of Significance of $R_{YY'}$

It is often useful to perform a significance test to see if the observed R_{YY} is significantly different from zero. In other words it is desired to test the null hypothesis H_0 $R_{YY'} = 0$ against the alternative hypothesis H_1 $R_{YY'} \neq 0$ If all the regression coefficients B are equal to zero, then SS_{reg} will be equal to zero and R_{YY} will also be equal to zero. The obtained values of $b'_i s$ are estimates of the corresponding population values β_i. A test of the null hypothesis that the all the population values are equal to zero will be given by

$$F_o = \frac{R_{YY}^2/k}{(1 - R_{YY'}^2)/(n - k - 1)} \tag{5 32}$$

where k the number of X variables If both the numerator and the denominator of eq ue multiplied by SS_{tot}, then the equivalent test,

$$F_o = \frac{SS_{reg}/k}{SS_{res}/(n-k-1)} = \frac{MS_{reg}}{MS_{res}} \tag{5 33}$$

For both F ratios the degrees of freedom for the numerator will be equal to k, the number of X variables, and the degrees of freedom for the denominator will be equal to $n\,k\,1$ where n is the number of samples This observed value is compared with the $F_{\alpha/2\ v1\ v2}$, with different α values Here the values of $v1$ and v are k and $n\,k\,1$

5 8 2   Test of Significance of B

For a test of significance of a regression coefficient in a multiple regression one can find the observed value of the test by using below equation

$$t_o = \frac{b_i}{\sqrt{\frac{MS}{\sum x^2(1-R^2)}}} \tag{5 34}$$

The square of t test will be equal to F test with 1 and $n\,k\,1$ degrees of freedom, so one can do F test for this and compare with the $F_{\alpha/2\ v1\ v2}$ In this case the values of $v1$ and v are 1 and $n\,k\,1$

5 9   Multicollinearity

The existence of high inter correlations among the X variables is called the prob lem of multicollinearity Evidence regarding a high degree of multicollinearity can be provided by the standard errors of the regression coefficients Form below equation

$$S_b = \sqrt{\frac{MS_{res}}{\sum x_i^2(1-R^2)}} \tag{5 35}$$

where R_i^2 is the squared multiple correlation of X_i with the remaining X vari ables As R_i^2 approaches 1, the denominator becomes small and S_b becomes large An extremely large value of S_b for any of the regression coefficients may be an indication of multicollinearity

5 10   Algorithm for Multiple Regression

The detailed steps that have to be done in *Multiple Regression Analysis* are given is this section   Multiple regression equation, as explained before is given below

$$Y' = a + b_1 X_1 + b_2 X_2 +, \quad , + b_k X_k \tag{5 36}$$

One have to find the regression coefficients $a, b_1, b_2, \quad , b_k$ by using the his torical measurements between independent variables (X) and dependent variable (Y) After that one have to do statistical analysis of the data and test of signif icance of the regression coefficients The procedure of all analysis is carried from Step 1 to Step 12

> **Step 1**   *Load the historical data of Independent $(X_{K \times N})$ and depen dent variables $(1 \times N)$, where N is the number of points in the training data bank and K is the number of Independent variables*

> **Step 2**   *Form the Intercorrelation (R_{ij}) and Correlation matrices (R_{ij}^{-1})*

> **Step 3**   *Find the inverse of Intercorrelation matrix that is R_{ij}^{-1}*

> **Step 4**   *Find B_i which is called standard form of regression coefficient vector by below equation*
> $$B_i = R_{ij}^{-1} R_{Yi} \tag{5 37}$$

> **Step 5**   *Now find the regression coefficient vector $(B = [b_1, b_2, \quad b_k])$ by using below equation*
> $$b_i = \frac{s_Y}{s_X} b_i \tag{5 38}$$
> *where $i = 1, 2, 3, \quad , k$, s_Y and s_X are the standard deviations of dependent and independent variables*

> **Step 6**   *Calculate intercept coefficient (a) by using below equation*
> $$a = \overline{Y} - b_1 \overline{X}_1, \quad , - b_k \overline{X}_k \tag{5 39}$$
> *where $\overline{X}_i$ and $\overline{Y}$ are the mean values of independent and dependent variables*

Step 7   *Calculate the multiple correlation coefficient ($R_{Y'Y}^2$) by below equation*

$$R_{Y'Y}^2 = [B_i]^T [R_{Yi}] \tag{5 40}$$

Step 8   *Now find all sum of squares which are denoted by SS_{tot}, SS_{reg} and SS_{res} by using below equations*

$$SS_{tot} = \sum y^2 \tag{5 41}$$

where $y = Y - \overline{Y}$

$$SS_{reg} = b_1 \sum x_1 y + b_2 \sum x_2 y + \quad + b_k \sum x_k y \tag{5 42}$$

where $x_i = X_i - \overline{X}$

$$SS_{res} = SS_{tot} - SS_{reg} \tag{5 43}$$

Step 9   *Find the Mean Squares of each sum of squares by using below equations*

$$MS_{tot} = \frac{SS_{tot}}{Stot_{df}} \tag{5 44}$$

$$MS_{reg} = \frac{SS_{reg}}{Sreg_{df}} \tag{5 45}$$

$$MS_{res} = \frac{SS_{res}}{Sres_{df}} \tag{5 46}$$

where df stands for degrees of freedom

The degrees of freedom for each sum of square is $N - 1$, K and $N - K - 1$

Step 10   *Find the standard error of all regression coefficients that is $B's$ by below equation*

$$S_b = \sqrt{\frac{MS_{res}}{\sum x_i^2 (1 - R_i^2)}} \tag{5 47}$$

where $R_i^2 = 1 - \frac{1}{r^{ii}}$, in which r^{ii} is the ith diagonal element in R_{ij}^{-1}

6 Artificial Neural Networks

6 1   Introduction

Artificial neural networks are parallel models comprised of densely interconn ected adaptive processing units These networks are fine grained parallel imple mentations of nonlinear static or dynamic systems A very important feature of these networks is their adaptive nature, where "learning by example" replaces "programming" in solving the problems This feature makes such computational models very appealing in application domains where one has little or incomplete understanding of the problem to be solved but where training data is readily available [10]

Artificial neural networks are viable computational models for a wide variety of problems These include pattern classification, forecasting and prediction, clus tering, nonlinear system modeling etc In this chapter it is mainly concerned on the "forecasting and prediction" type application by using the well known *Error Back Propagation (EBP)* algorithm

This chapter is organized mainly in five parts in which, section 6 2 discusses the basic model of the artificial neural networks In section 6 3 consists about the learning process in the neural network concentrating mainly on *Gradient Descent Based learning rule* In section 6 4 detailed description of *EBP* algo rithm is explained, where as section 6 5 points the modifications of *EBP* and its limitations
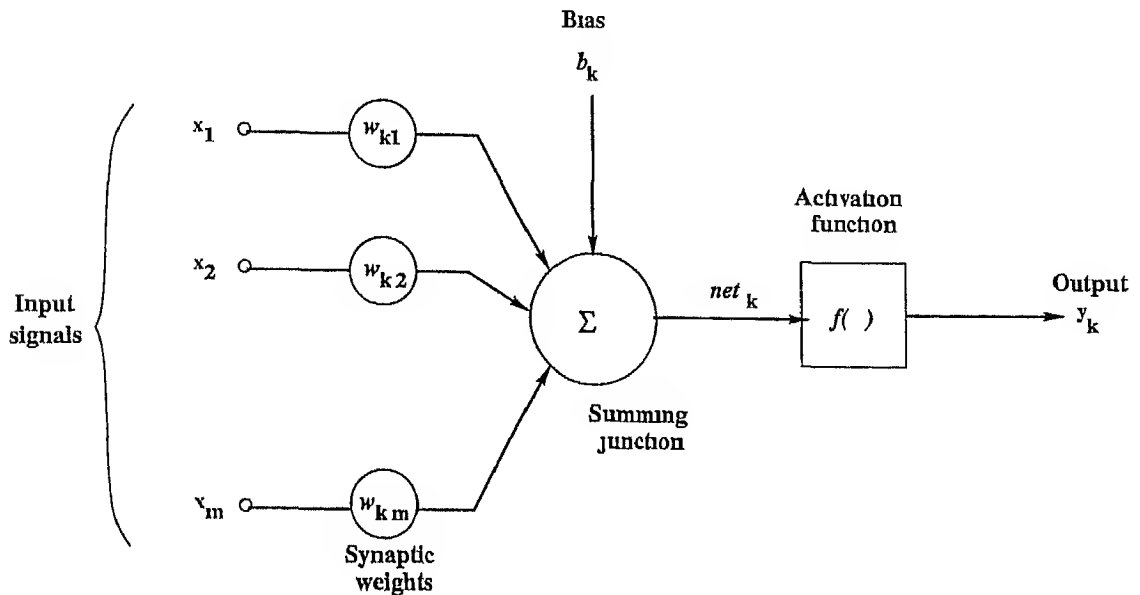
Bias

b_k

Activation
function

x_1 w_{k1}

x_2 w_{k2}

Input
signals

Σ net_k $f(\ )$ Output
y_k

Summing
junction

x_m w_{km}

Synaptic
weights

Figuie 6 1   The basic model of a neuron

6 2   The Basic Model of Artificial Neural Network

The fund imcntal to any neuial network is it's information processing unit, which is called is *ncuion* The block diagram of figuie 6 1 shows the model of a neuron, which foims the basis for designing artificial neural networks [11] Every neuron consists thiee basic elements which are called synaptic connections, summing point and an activation function, which aie shown in figuie 6 2

Synaptic connections Synaptic connections are mainly the connections between input signal (x_i) of the neuion to the summing point It is characterized by i weight (w) or strength of its own As an example a connection weight, between ith input signal (x_i), to the jth summing point can be denoted by w_{ji}

Summing point It is the point wheie all the input signals multiplied by their synaptic weights will be added

Activation function This is for limiting the amplitude of the output (y_k) of a neuion The activation function is also iefeiied to as a squashing function

in that it squashes (limits) the permissible amplitude range of the output signal to some finite value. Typically, the normalized amplitude range of the output of a neuron is written as the closed unit interval [0 1] or alternatively [1,1]

The neuronal model of figure also includes an externally applied connection which is called a bias. This is denoted in figure as b_k. The bias has the effect of increasing or lowering the net input of the activation function, depending on whether it is positive or negative, respectively. In mathematical terms, neuron k can be described by writing the following equations

$$net_k = \sum_{i=1}^{m} w_{ki} x_i + b_k \qquad (6.1)$$

and

$$y_k = f(net_k) \qquad (6.2)$$

where $x_1, x_2, \quad x_m$ are the input signals, $w_{k1}, w_{k2}, \quad w_{km}$ are the synaptic weights of neuron k, net_k is the linear combiner output due to the input signals and bias, b_k is the bias signal, $f(\ )$ is the activation function, and y_k is the output signal of the neuron

6 2 1   Types of Activation Function

Mostly in neuron models choosing the activation function is important. Here various activation functions which are found in literature [11] will be explained

Threshold Function   The graphical form of this function is shown in figure 6 2 (a) The mathematical representation of this function is

$$f(net) = \begin{cases} 1 & if \quad net \geq 0 \\ 0 & if \quad net < 0 \end{cases} \qquad (6.3)$$

Such a neuron is referred to in the literature as the *McCulloch Pitts model*. In this model, the output of a neuron takes on the value of 1 if the *net* of the neuron is nonnegative, and 0 otherwise
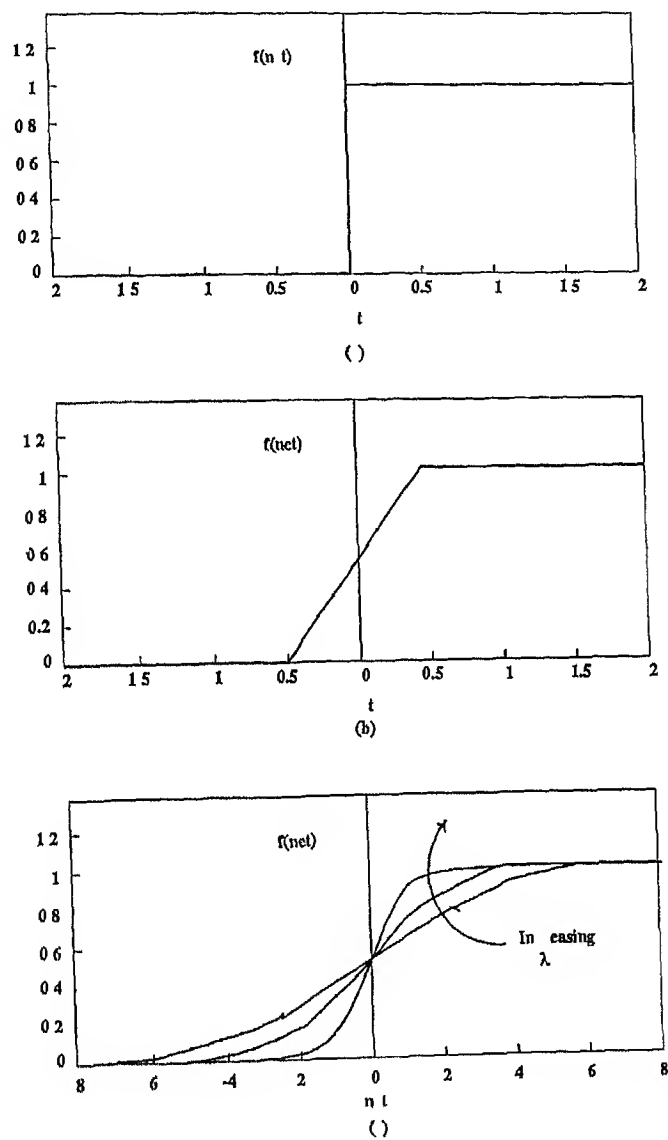
Figure 6 2   Different activation functions (a) Threshold function (b) Piecewise
linear function (c) Sigmoid function

Piecewise Linear Function   This is shown in figure 6 2 (b) , the mathematical representation of this function is

$$f(net) = \begin{cases} 1 & \text{if net} \geq +1/2 \\ net & \text{if net} < +1/2 > net > -1/2 \\ 0 & \text{if net} \leq -1/2 \end{cases} \tag{6 4}$$

where the amplification factor inside the linear region of operation is assumed to be unity This form of an activation function may be viewed as an *approximation* to a nonlinear amplifier

Sigmoid Function   The sigmoid function, whose graph is s-shaped, is by far the most common form of activation function used in the construction of artificial neural networks It is defined as a strictly increasing function that exhibits a balance between linear and nonlinear behavior An example of the sigmoidal function is the *logistic function*, difined by

$$f(net) = \frac{1}{1 + exp(-\lambda net)} \tag{6 5}$$

where $\lambda > 0$ is called the steepness parameter of the sigmoidal function It's graphical view is shown in figure 6 2 (c) for different values of λ The main feature of this type of activation function is, it is differentiable, where as threshold function is not In Error Back Propagation which is discussed later in this chapter, is used maily sigmoidal function, which uses its differentiable function

The bias b_k is an external parameter of artificial neuron k, can be formulted as a another input signal x_0, which has a value of ± 1, connected with a synaptic weight of value b Now the model of a k th neuron can be written by following equations

$$net_k = \sum_{i=0}^{m} w_{ki} x \quad \text{and}$$

$$y_k = f(net_k) \tag{6 6}$$

where x is input signal vector, having $x_0 = \pm 1$, w is the weight vector having $w_{k0} = b$

As explained before that neural networks are densily interconnected neurons it is possible to get different types of neural network structure by choosing different neuron models with different activation functions   One of such type of network is *Backpropagation neural network* which is explained later   The other main important thing in neural network is "learning algorithm"   It is explained in comming section

6 3   Learning Rule

One of the most significant attributes of a neural network is its ability to learn by interacting with its environment or with an information source   Learning in a neural network is normally accomplished through an adaptive procedure, known as a learning rule or algorithm, whereby the weights of the network are incre mentally adjusted so as to improve a predefined performance measure over time The learning process can be viewed as "search" in a multidimensional parameter (weight) space for a solution, which gradually optimizes a prespecified objective (criterion) function

There are number of learning rules presented in the literature [10] which are basically classified into three types of learning tasks   They are supervised, un supervised and reinforced learning tasks   In supervised learning (also known as learning with a teacher), each input pattern received from the environment is associated with a specific desired target pattern   Usually, the weights are syn thesized gradually, and at each step of the learning process they are undated so that the error between the network's output and a corresponding desired target is reduced   On the other hand, unsupervised learning (known as learning with out teacher) involves the clustering of unlabeled patterns of a given training set Reinforcement learning involves updating the network s weights in response to an "evaluative" teacher signal, this differs from supervised learning   where the teacher signal is the *correct answer*

The following general learning rule is adopted in neural network studies [28]   The weight vector $w_i = [w_{i1}, w_{i2}, \quad , w_{in}]^t$ increases in proportion to the product
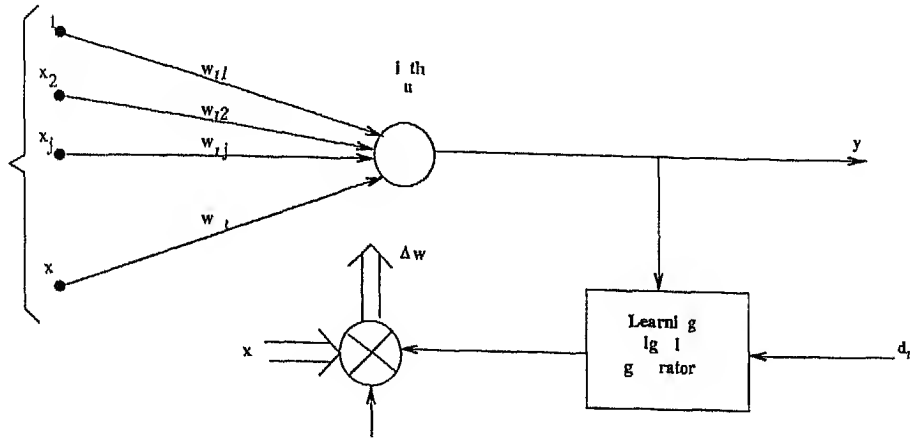
Figure 6 3   Illustration for weight learning rules

of input x and learning signal r The learning signal r is in general a function
of w_i, x, and sometimes of the teacher's signal d_i The whole thing is shown in
figure 6 3

$$r = r(w_i, x, d_i) \tag{6 7}$$

The increment of weight vector w_i produced by the learning step at time t ac
cording to the general learning rule is

$$\Delta w_i(t) = cr[w_i(t), x(t), d_i(t)]x(t) \tag{6 8}$$

where c is a positive number called the learning constant that determines the rate
of learning   The weight vector adapted at time t becomes at the next instant, or
learning step,

$$w_i(t+1) = w_i(t) + cr[w_i(t), x(t), d_i(t)]x(t) \tag{6 9}$$

This can be written in formal notations

$$w_i^{k+1} = w_i^k + cr(w_i^k, x^k\, d_i^k)x^k \tag{6 10}$$

where w_i^k is the weight vector at k th step and w_i^{k+1} is the k+1 th iteration
weight vector
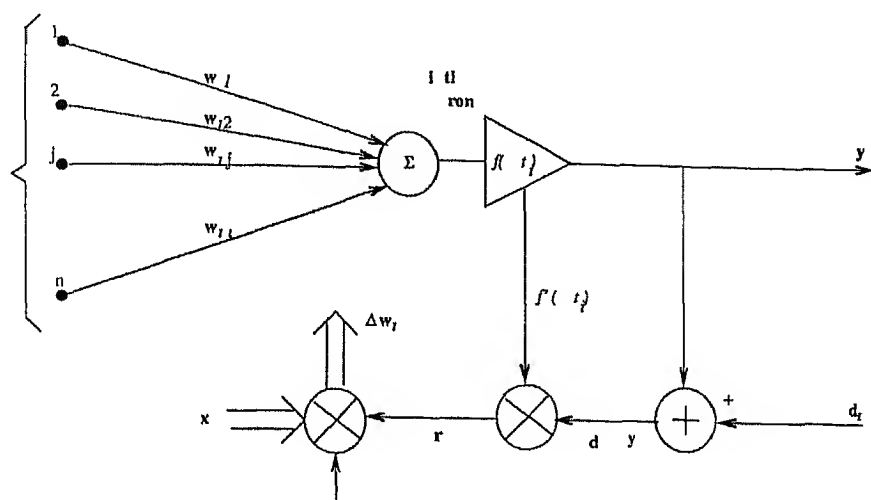
Figure 6 4   Delta Learning Rule

In [28], various learning algorithms are discussed, based on this basic learning algorithm Among them "Gradient Descent based Delta Algorithm" is one which gives basic idea about the *Error Back Propagation Algorithm* The main aim of this chapter is to explain the EBP, so it is better to know the "Delta   Rule", which is explained later in this section

6 3 1   Delta-Rule

The delta learning rule is only valid for continuous activation functions like figure 6 2 (c) The learning signal for this is called delta and is defined as [28]

$$r \triangleq [d_i - f(net)]f'(net) \tag{6 11}$$

the term $f'(net)$ is the derivative of the activation function $f(net)$ computed for $net = w_i^t x$ The proceeding of the delta learning rule is indicated in figure 6 4 This learning rule can be readily derived from the condition of least squarred error between y_i and d_i The sum of squarred error defined as

$$E \triangleq \frac{(d_i - y_i)}{2} \tag{6 12}$$

where $y_i = f(w_i^t a)$ The gradient of this error is given by

$$\nabla E = -(d_i - y_i) f'(w_i^t x) a \tag{6 13}$$

The components of gradient vector are

$$\frac{\partial E}{\partial w_{ij}} = -(d_i - y_i) f'(w_i^t x) x \quad \text{for} \quad j = 1, 2, \quad , n \tag{6 14}$$

Since the minimization of the error requires the weight changes (ΔW) to be in the negative gradient direction, so

$$\Delta w_i = -\eta \nabla E \tag{6 15}$$

where η is a positive constant known as learning rate Then one can obtain from equation 6 13 and equation 6 15

$$\Delta w_{ij} = \eta(d_i - y_i) f'(net_i) x_j \quad \text{for} \quad j = 1, 2, \quad , n \tag{6 16}$$

One disadvantage of the delta learning rule is , it progress very slowly when the f'(net)=0, even when the error is large f'(net) = 0 when net has large magnitude To avoid this the common technique is replace f' by f' plus a small positive bias ϵ [10] In this case the weight adjustment becomes

$$\Delta w_{ij} = \eta(d_i - y_i)(f'(net_i) + \epsilon) x_j \quad \text{for} \quad j = 1, 2, \quad , n \tag{6 17}$$

one of the primary advantages of the delta rule is that it has a natural extension that may be used to train multilayered neural nets

6 4   Back Propagation

The back propagation algorithm is a supervised learning method for multi layered feedforward neural networks It is essentially a gradient descent local optimization technique, it involves backward error correction of the network weights The backpropagation learning rule is central to much current work on learning in artificial neural networks Back propagation trained multilayer neural nets have been applied successfully to solve some difficult and diverse problems, such as time series prediction, non linear system modelling, pattern classification etc [10]
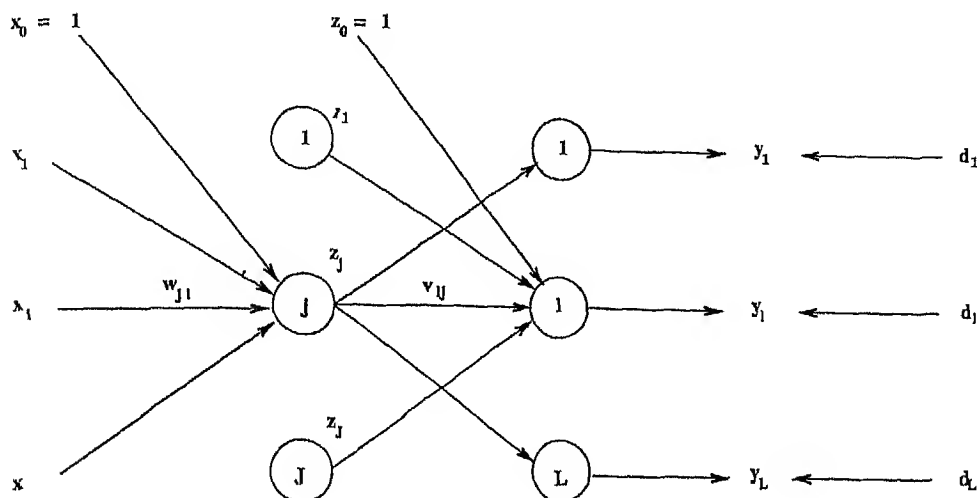
Figure 6 5   A two layer fully interconnected feedforward neural network architecture

6 4 1   Learning Rule for Multilayer Feedforward Neural Networks

Consider the two layer feedforward architecture shown in figure 6 5   This network receives a set of scalar signals $x_0, x_1,$ $, x_n$ where x_0 is a bias signal equal to 1   This set of signals constitutes an input vector $x \in R^{n+1}$ The layer receiving the input signal is called the hidden layer   Figure 6 5 shows a hidden layer having J units   The output of the hidden layer is a $(J+1)$ dimensional real valued vector $z = [z_0, z_1,$ $, z_J]$ Again, $z_0 = -1$ represents a bias input   The vector z supplies the input for the output layer of L units   The output layer generates an L dimensional vector y in response to the input x which, when the network is fully trained, should be identical (or very close) to a desired output vector d associated with x

The activation function f_h of the hidden units is assumed to be a differentiable nonlinear function (typically, f_h is the logistic function defined by $f_h(net) = 1/(1 + e^{-\lambda net})$ with values of λ close to unity)   Each unit of the output layer is assumed to have the same activation function f_o The functional form of f_o is determined by the desired output signal/pattern representation or the type of application   For example, if the application is of *estimation type* with desired

values real positive, then a saturating nonlinearity similar to f_h may be used for f_o. Finally denote w_{ji} the weight of the j th hidden unit associated with the input signal x_i. Similarly v_{lj} is the weight of the l th output unit associated with the hidden signal z_j.

Next, consider a set of m input/output pairs x_k, d_k, where d_k is an L dimensional vector representing the desired network output upon presentation of x_k. The object here is to adaptively adjust the w, v weights of this network such that the underlying function/mapping represented by the training set is approximated or learned. Since the learning here is supervised, and error function may be defined to measure the degree of approximation for any given setting of the network's weights. A commonly used error function is the sum of squared error (SSE) measure. Once a suitable error is formulated, learning can be viewed as an optimization process. That is, the error function serves as a criterion function, and the learning algorithm seeks to minimize the criterion function over the space of possible weight settings. If a differentiable criterion function is used, agradient descent on such a function will naturally lead to a learning rule. Now consider the following error function which is minimized by adjusting the weights

$$E(w) = \frac{1}{2} \sum_{l=1}^{L} (d_l - y_l)^2 \tag{6 18}$$

Here w represents the set of all weights in the network.

6 4 2   Error Backpropagation Learning Rule

Since the targets for the output units are explicitly specified, one can use the delta rule directly for updating the v_{lj} weights. That is

$$\Delta v_{lj} = v_{lj}^{new} - v_{lj}^{c} = -\eta \frac{\partial E}{\partial v_{lj}} = \eta \delta_l z_j \tag{6 19}$$

$$where \quad \delta_l = (d_l - y_l) f_o'(net_l) \tag{6 20}$$

for $l = 1, 2, \quad L$ and $j = 0, 1, 3 \quad J$

Here $net_l = \sum_{j=0}^{J} v_{lj} z_j$ is the weighted sum of the l th output unit, f_o' is the derivative of f_o with respect to net, δ_l is the error signal of the l th output

unit and v_{ij}^{new} and v_{ij}^{c} represents the updated (new) and current weight values, respectively. The z_j values are computed by propagating the input vector x through the hidden layer according to

$$z_j = f_h\left(\sum_{i=0}^{n} w_{ji}x_i\right) = f_h(net_j) \quad j = 1,2, \quad J \qquad (6\ 21)$$

The learning rule for the hidden layer weight w_{ji} is not as obvious as that for the output layer because a set of target values for hidden units are not available. However, one may derive a learning rule for hidden units by attempting to minimize the output layer error. This amounts to propagating the output errors back through the output layer toward the hidden units in an attempt to estimate *dynamic targets* for these units. Such a learning rule is termed *error backpropagation* or the *backpropagation learning rule* and may be viewed as an extension of the delta rule.

To complete the derivation of backprop for the hidden layer weights, and similar to the preceding derivation for the output layer weights, gradient descent is performed on the criterion function in equation 6 18. This time the gradient is calculated with respect to the hidden weights

$$\Delta w_{ji} = -\eta\frac{\partial E}{\partial w_{ji}} \quad where \quad j = 1,2,3, \quad J, i = 0,1,2, \quad n \quad (6\ 22)$$

Using the chain rule for differentiation, one may express the partial derivative in equation 6 22 as

$$\frac{\partial E}{\partial w_{ji}} = \frac{\partial E}{\partial z_j}\frac{\partial z_j}{\partial net_j}\frac{\partial net_j}{\partial w_{ji}} \qquad (6\ 23)$$

with

$$\frac{\partial net_j}{\partial w_{ji}} = x_i, \tag{6 24}$$

$$\frac{\partial z_j}{\partial net_j} = f'_h(net_j), \tag{6 25}$$

and

$$\begin{aligned}
\frac{\partial E}{\partial z_j} &= \frac{\partial}{\partial z_j}\{\frac{1}{2}\sum_{i=1}^{L}[d_l - f_o(net_l)]^2\}\\
&= -\sum_{i=0}^{L}[d_l - f - o(net_l)]\frac{\partial f_o(net - l)}{\partial z_j}\\
&= -\sum_{i=0}^{L}(d_l - y_l)f'_o(net_l)w_{lj}\\
&= -\sum_{i=0}^{L}\delta_l w_{lj}
\end{aligned} \tag{6 26}$$

combining equation 6 25 and equation 6 26 then

$$\delta_j = f'_h(net_j)\sum_{i=0}^{L}\delta_l w_{lj} \tag{6 27}$$

where δ_j is the error signal of j th hidden neuron   The weight adjustment in the hidden layer becomes

$$\Delta w_{ji} = \eta \delta_j x_i \tag{6 28}$$

It is usually possible to express the derivatives of the activation functions in equation 6 19 and equation 6 27 in terms of the activations themselves   For example, for the logistic activation function, $f(net) = 1/(1 + e^{-\lambda net})$

$$f'(net) = \lambda f(net)[1 - f(net)] \tag{6 29}$$

and for the hyperbolic tangent activation function $f(net) = tanh(\beta net)$

$$f'(net) = \beta[1 - f^2(net)] \tag{6 30}$$

These learning equations may also be extended to feedforward nets with more than one hidden layer and/or nets with connections that jump over one or more layers

6 4 3   Error Back-Propagation Training

Figure 6 6 illustrates the flowchart of the error back propagation training algo rithm for a basic two layer network as in figure 6 5 The learning begins with the feedforward recall phase (Step 2) After a single pattern vector x is submitted at the input, the layers responses x and y are computed in this phase Then, the error signal computation phase (Step 4) follows Note that the error signal vector must be determined in the output layer first, and then it is propagated toward the network input nodes The $L \times J$ weights are subsequently adjusted within the matrix V in Step 5 Finally $J \times N$ weights are adjusted within the matrix W in Step 6 Note that the cumulative cycle of input to output mapping is computed in Step 3 as a sum over all continuous output errors in the entire training set The final error value for the entire training cycle is calculated after each completed pass through the training set $x_1, x_2,$ x_p The learning proce dure stops when the final error value below the upper bound, E_{max}, is obtained as shown in Step 8 [28]

Figure 6 7 depicts the block diagram of the error back propagation trained net work operation and shows both the flow of signal, and the flow of error within the network [28] The feedforward phase is self explanator The shaded portion of the diagram refers to the feedforward recall The blank portion of the diagram refers to the training mode of the network The back propagation of error $d - y$ from each output, for $l = 1, 2, 3,$ L, using the negative gradient descent tech nique is divided into functional steps such as calculation of the error signal vector δ_l and calculation of the weight matrix adjustment ΔV of the output layer The diagram also illustrates the calculation of internal error signal vector δ_j and of the resulting weight adjustment ΔW of the input layer

The complete procedure for updating the weights in a feedforward neural net is summarized below for the two layer architecture of figure 6 5

Given there are P training patterns

Step 1   *Set the learning rate η to a small positive value Initialize*
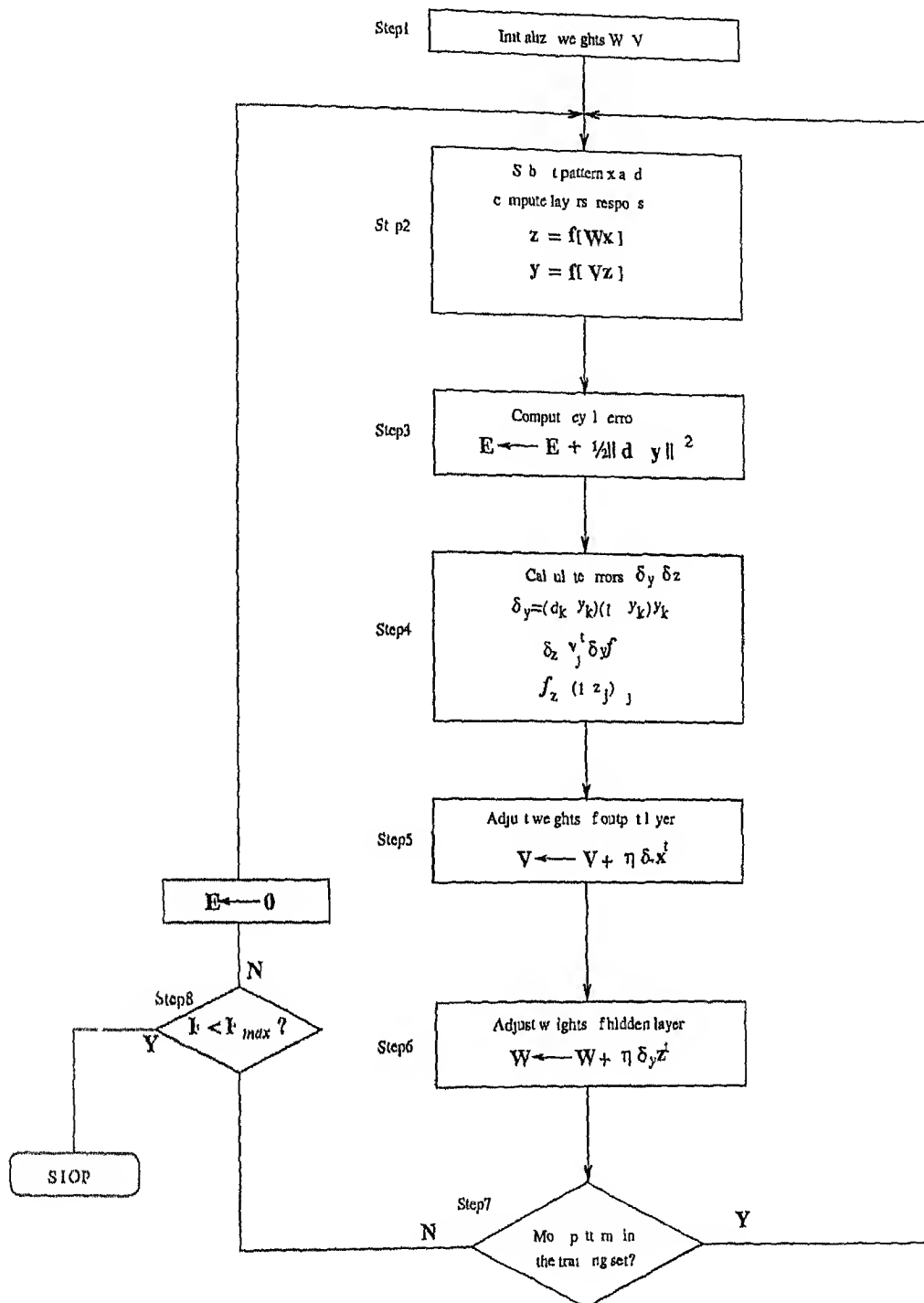   *all weights at small random values and refer them as current*

Figure 6 6   Error Back Propagation training ($EBPT$) algorithm flowchart
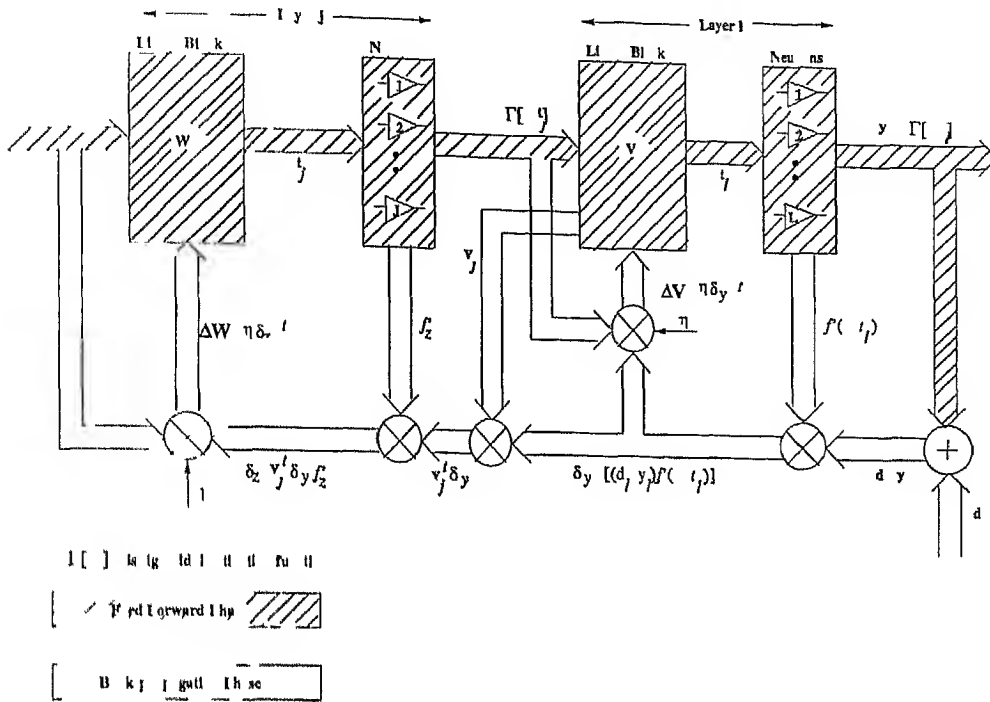
Figure 6 7   Error Back Propagation training (EBPT) algorithm block diagram

weights Choose E_{max}, $p = 1, E = 0$

Step 2   Select an input pattern x_k from the training set (preferably at random) and propagate it through the network, thus generating hidden and output unit activities based on the current weight settings

Step 3   Compute the error $E = E + \frac{1}{2}(d_k - y_k)^2$

Step 4   Compute the error signals δ_l and δ_j by employing the equations 6 20 and 6 27

Step 5   Adjust output layer weights V as shown in equation

$$v_{lj}^{new} = v_{lj}^{old} + \Delta v_{lj} \tag{6 31}$$

Step 6   Adjust hidden layer weights W as shown in equation

$$w_{ji}^{new} = w_{ji}^{old} + \Delta w_{ji} \tag{6 32}$$

Step 7   if $p < P$ then $p = p + 1$, and goto step 2   other wise goto step 8

Step 8   One training cycle is completed   Now test for convergence
If $E < E_{max}$ then terminate the training session, else put $E = 0$ and $p = 1$ initialize the new training cycle by going to step 2

It should be possible that backpropagation may fail to find a solution that passes the convergence test   In this case, one may try to reinitialize the search process tune the learning parameters, and/or use more hidden units   The above procedure is called *incremental backprop learning* because the weights are updated after every presentation of an input pattern   This incremental updating is more desirable due to two reasons   (1) it requires less storage ,and (2) it makes the search path in the weight space stochastic (here, at each time step, the input vector x is drawn at random), which allows for a wider exploration of the search space and, potentially leads to better quality solution

6 5   Backpropagation Enhancements and Variations

In general, learning with backpropagation is slow   Typically, this is due to the characteristics of the error surface [10]   The surface is characterized by numerous flat and steep regions   In addition, it has many troughs that are   flat in the direction of search   Sometimes when backpropagation converges, it may converge to a local minimum of the criterion function   This fact is true in any gradient-descent based learning rule when the surface is being searched in nonconvex, i e , it admits local minima   The local minima problem can be explained as shown in figure 6 8   Suppose one started with a weight set for the network corresponding to point P   Due to gradient descent learning, the minimum one can encounter is the one at M_1 not that at M_g M_1 is called a *local minimum* and corresponds to a partial solution for the network in response to the training data M_g is the global minimum point which one have to get, unless measures
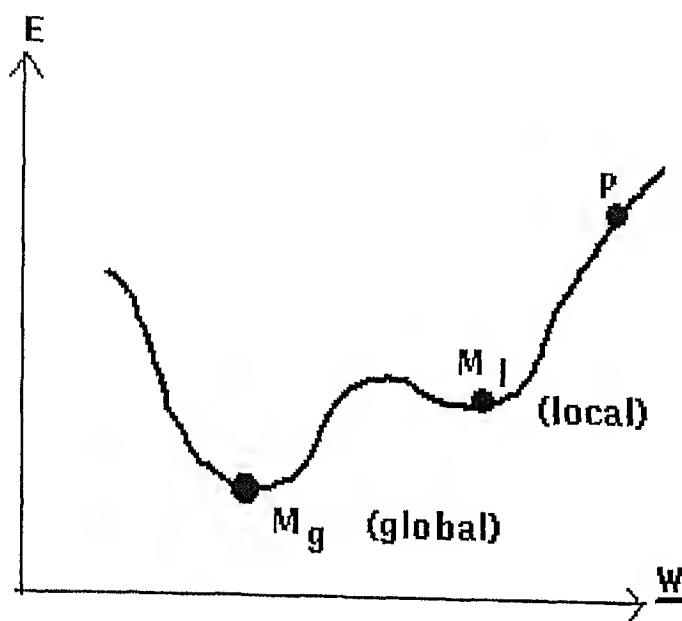
Figure 6 8   Local Minimum

are taken to escape from M_1, M_g will never be reached   Many enhancements of and variations to backprop have been proposed   These are mostly heuristic modifications with goals of increased speed of convergence, avoidance of local minima, and/or improvement in the network's ability to generalize   Some of them are presented here which are mainly considered by the author in his project

6 5 1   Weight Initialization

Owing to its gradient descent nature, backpropagation is very sensitive to initial conditions   If the choice of the initial weight vector w^o happens to be located nearer to the solution point, then the convergence of back prop is fast   On the other hand, backprop converges very slowly if w^o starts the search in a relatively flat region of the error surface   In practice, the weights are normally initialized to small zero mean random values   The motivation for starting from small weights is that large weights tend to prematurely saturate units in a network and render them insensitive to the learning process   On the other hand, randomness is

introduced as a symmetry breaking mechanism it prevents units from adopting similar functions and becomming redundant [10]

6 5 2   Learning Rate

The convergence speed of backprop is directly related to the learning rate parameter η If η is small, the search path will closely approximate the gradient path, but convergence will be very slow due to the large number of update steps needed to reach a minima   On the other hand if η is large convergence initially will be very fast, but the algorithm will eventually oscillate and thus not reach a minimum   In general, it is desirable to have large steps when the search point is far away from a minimum, with decreasing step size as the search approaches a minimum   Many heuristics have been proposed so as to adapt the learning rate automatically   In [10] used a method where the learning rate for a given weight w_i is to set to $a\eta_i(t)$ if $\frac{\partial E(t)}{\partial w}$ and $\frac{\partial E(t-1)}{\partial w}$ have the same sign, with $a > 1$ if the partial derivatives have different signs, then a learning rate of $b\eta_i(t)$ is used, with $0 < b < 1$ A similar, theoretically justified method for increasing the convergence speed of incremental gradient descent search is to set $\eta(t) = \eta(t-1)$ if $\nabla E(t)$ has the same sign as $\nabla E(t-1)$, and $\eta(t) = \eta(t-1)/2$ otherwise

6 5 3   Momentum

Another simple approach to speed up backpropagation is through the addition of a momentum term to the right hand side of the weight update rules in equations 6 19 and 6 28   Here, each weight change Δw_i is given some momentum so that it accelerates in the average downhill direction instead of fluctuating with every change in the sign of the associated partial derivative $\frac{\partial E}{\partial w}$ The addition of momentum to gradient search is stated formally as

$$\Delta w_i = -\eta \frac{\partial E}{\partial w_i(t)} + \beta \Delta w_i(t-1) \tag{6 33}$$

where β is a momentum rate normally chosen between [0,1] and $\Delta w_i(t-1) = w_i(t) - w_i(t-1)$ Equation 6 33 is a special case of multistage gradient methods that have been proposed for accelerating convergence and escaping local minima
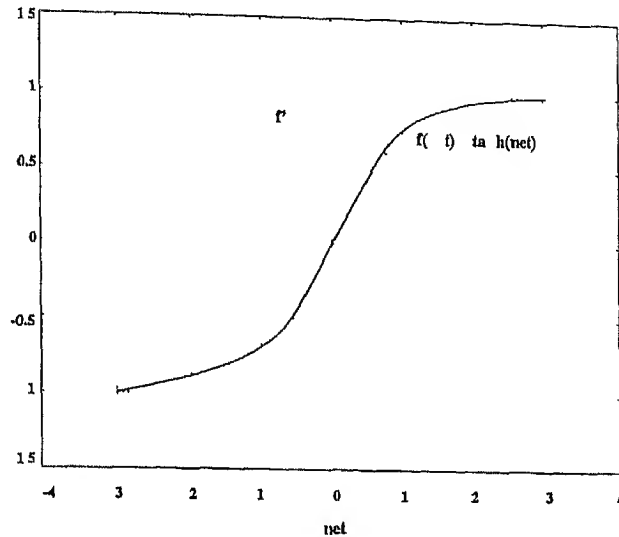
Figure 6 9   Plots of f(net) = tanh(net) and it's derivative f'(net)

6 5 4   The Flat Spot Problem

During the training, if a unit in a multilayer network receives a weighted signal *net* with a large magnitue, this unit outputs a value close to one of the saturation levels of its activation function If the corresponding target value is substantially different from that of the saturated unit, one can say that the unit is incorrectly saturated or has entered a flat spot When this happens, the size of the weight update due to backprop will be very small, even though the error is relatively large, and it will take an excessively long time for such incorrectly saturated units to reverse their states This situation can be explained by referring to figure 6 9 where the activation function $f(net) = tanh(\beta net)$ and its derivative $f'(net) = \beta[1 - f^2(net)]$ are plotted for $\beta = 1$ Here, when (net) has large magnitude, f' approaches zero Thus the weight change approaches zero in equations 6 19 and 6 28 even when there is difference between the actual and desired output for a given unit A simple solution to the flat spot problem is to bias the derivative of the activation function, that is, replace f'_o and f'_h in equations 6 20 and 6 27 by $f'_o + \epsilon$ and $f'_h + \epsilon$ respectively (a typical value for ϵ is 0 1)

7 Conclusions

This thesis was focused on one of the problems concerning the trading of renew able energies, that is, on the fluctuations of the renewable production. Those fluctuation problems apply specially to wind and photovoltaic power production. A marketer who buys such "fluctuating renewable energy", could possible order energy from conventional power markets to compensate the fluctuation of the renewable production. To know how much he should order, he would need a good forecast for the renewable production in the comming hours/days. Within the scope of this work different power forecast models for wind and photovoltaic energy were established and compued.

Two wind power forecast models were made, a fuzzy based model and a multiple regression based model. The results from the fuzzy model were acceptable. With some advancement the model would probably give a better forecast. More of the influencing factors could for example be taking into account and a better fuzzy rule base could be built. The regression model gave considerable better results than the fuzzy model. But also here some advancement could be made. In [21] it is mentioned, that the wind power has relationship to the cube of the wind speed. This could be taken into account in the regression model in a hope for better results. And, like for the other model, some improvement could possibly be reached through higher number of input variables.

For the photovoltaic power forecasting two models were constructed, a neural network based model and a multiple regression based model. The results from both the models were quite good. The neural network model gave slightly better results than the regression model. The used data included recognizable noise and it would be interesting to test the models with more accurate data.

In this thesis it has be shown that with suitable models relatively good renewable power forecasts can be made Such models could be extremely helpful for a power marketer, who is concentrated on buying and selling renewable energy With the help of such models the renewable energies could possibly be able to play more important role on the power markets in the future, than they have done until now

A. Additional Figures and Tables

A 1   Wind Power Forecasting

S No	Wind Speed m/s	Actual Power kW	Estimated Power kW	Error
1	3 2	48 6	45 013560	0 07379493
2	4 7	63 5	117 536520	0 850968
3	2 7	34 4	36 226521	0 0530965
4	5 5	69 3	156 288002	1 255238
5	6 1	104 9	175 042783	0 668663
6	4 6	90 2	113 456706	0 257834
7	6 3	58 5	185 8826813	2 177481
8	3 8	52 7	67 1450844	0 2741002
9	2 4	20 9	35 595315	0 7031251
10	4 9	124 2	142 250257	0 14533

Table A 1   Wind power forecasting results using fuzzy model with regularity information "regular"
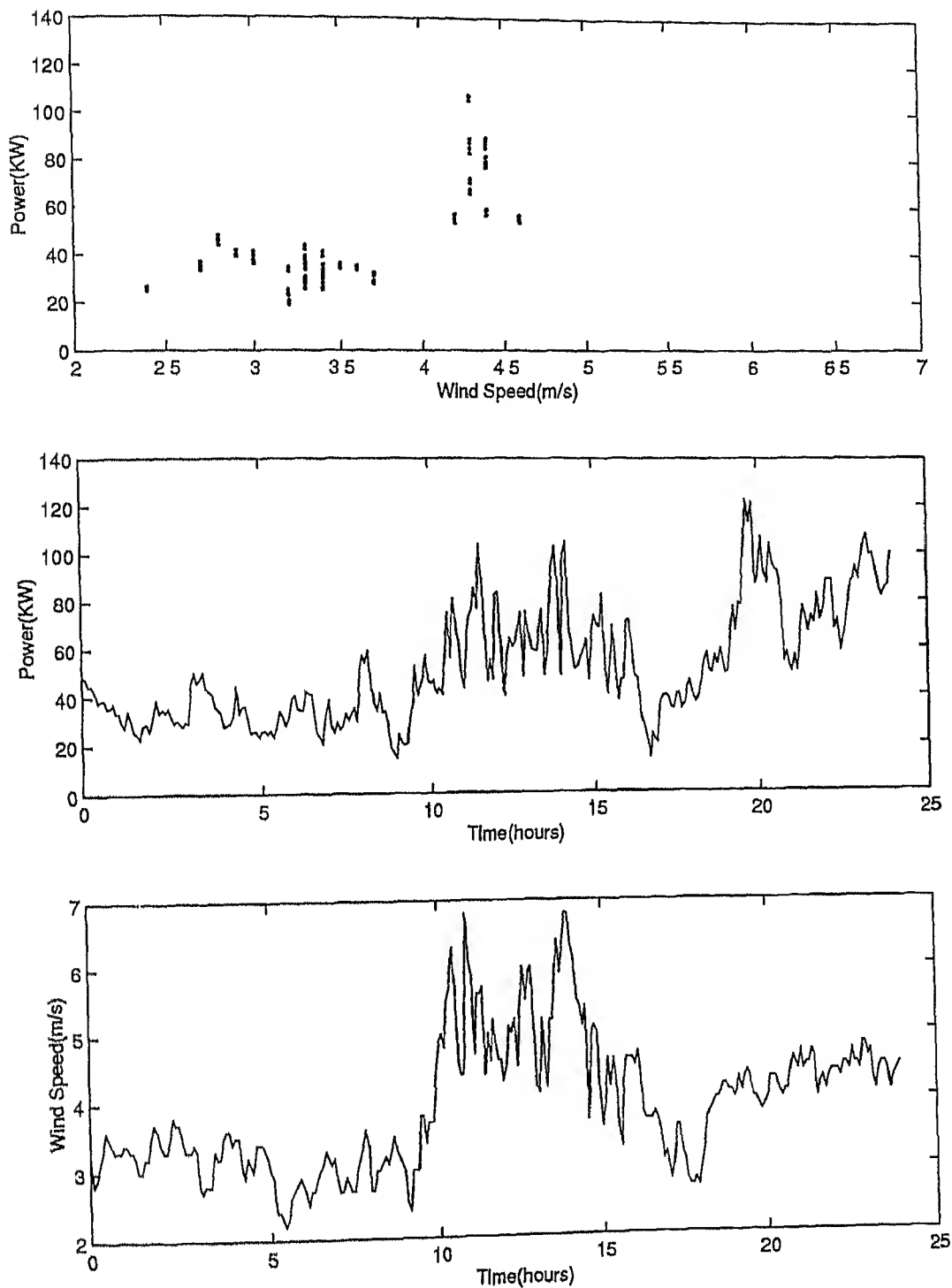
Figure A 1   18th October Oster wind plant data
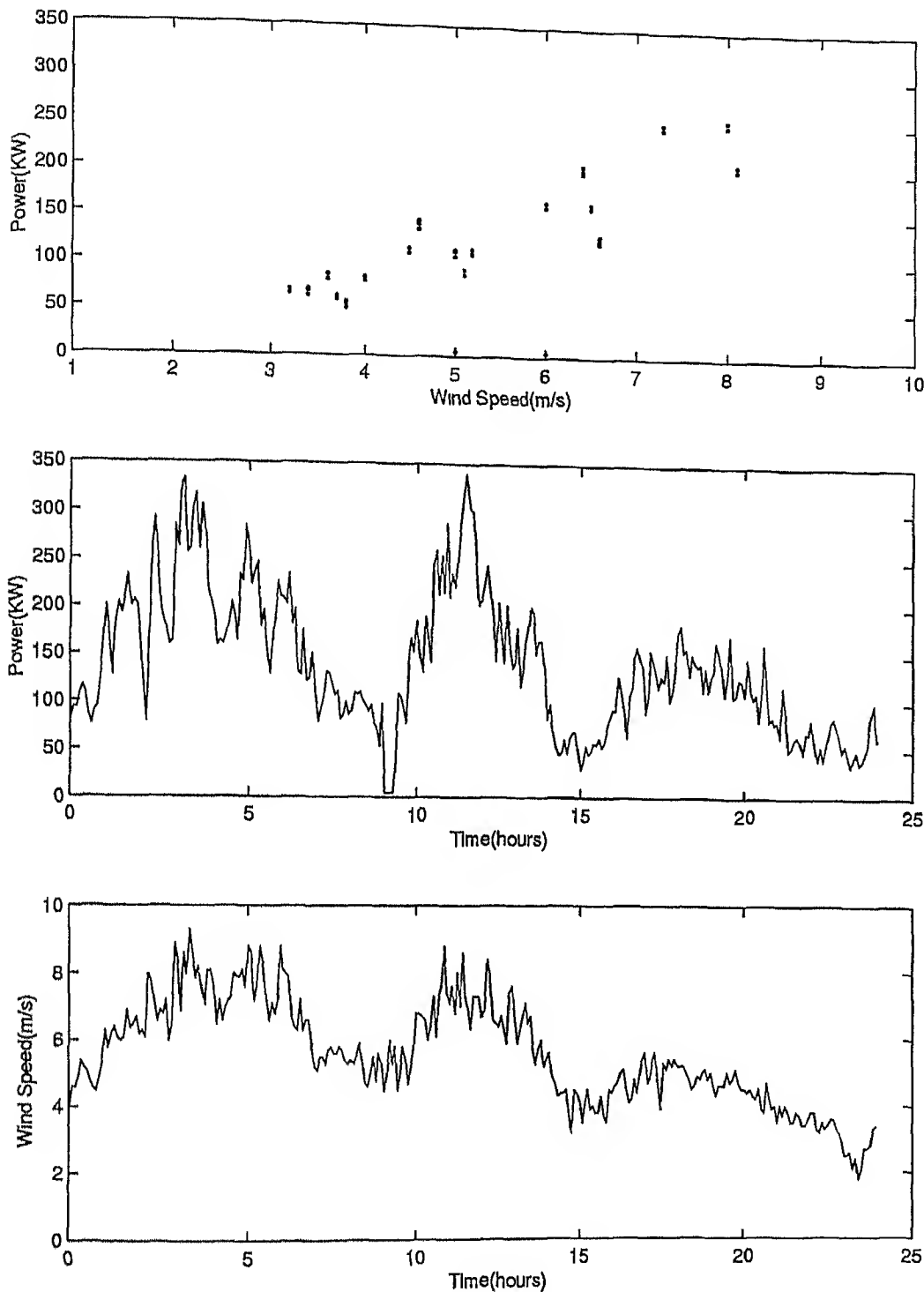
Figure A 2   14th November Oster wind plant data
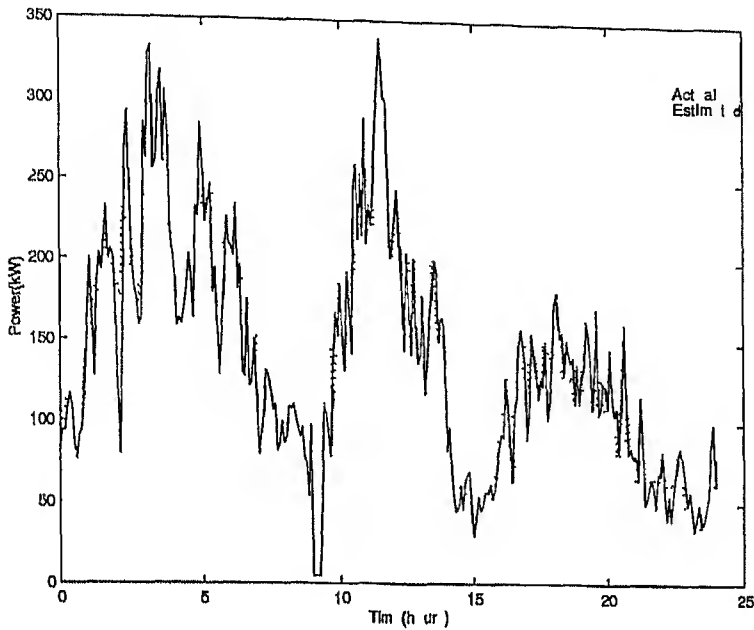
Figure A 3   14th November fuzzy based model result with regularity information
"little regular"
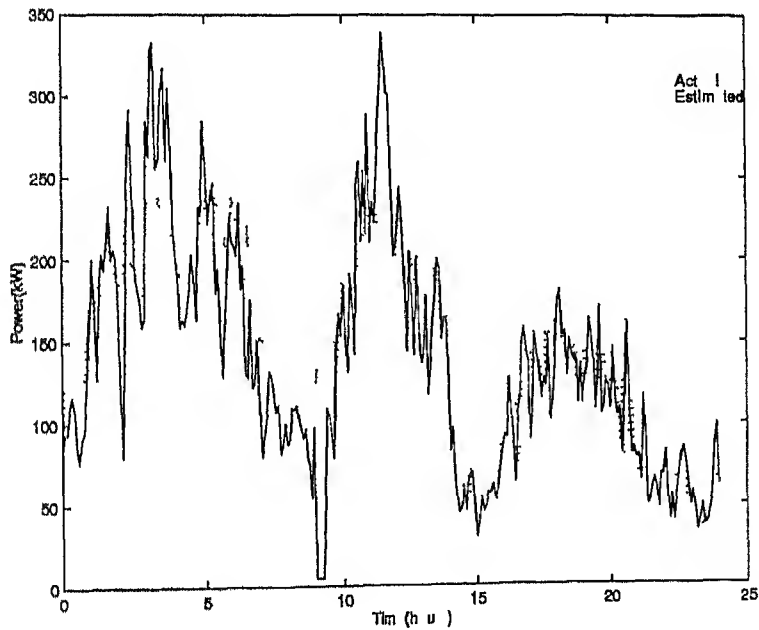
Figure A 4   14th November fuzzy based model result with regularity information
"very regular"

A 2   Photovoltaic Power Forecasting

Variable	b	S_b	F_b	Correlation with Power
Radiation	0 8788	0 01077	6651 304	0 96943
Temperature	0 1184	0 011781	101 0244	0 68491
Wind Speed	0 05608	0 01329	17 8029	0 59891

$R^2_{I\,(T\,Ra)} = 0\,0058$	$R_{P\,(RaT)} = 0\,972415$	$R^2_{P(W\,RaT)} = 9\,376\text{E-}004$
$F_{R^2} = 5994\,937$		

Table A 2   20th June data multiple regression analysis

Variable	b	S_b	F_b	Correlation with Power
Radiation	0 91547	0 0444	424 994	0 81403
Temperature	0 12543	0 03627	11 9556	0 6529
Wind Speed	0 00863	0 0268	0 10376	0 26614

$R^2_{P(T\,Ra)} = 0\,0066$	$R_{P\,(RaT)} = 0\,8181$	$R^2_{P(W\,RaT)} = 1\,963\text{E-}004$
$F_{R^2} = 402\,4229$		

Table A 3   9th January data multiple regression analysis
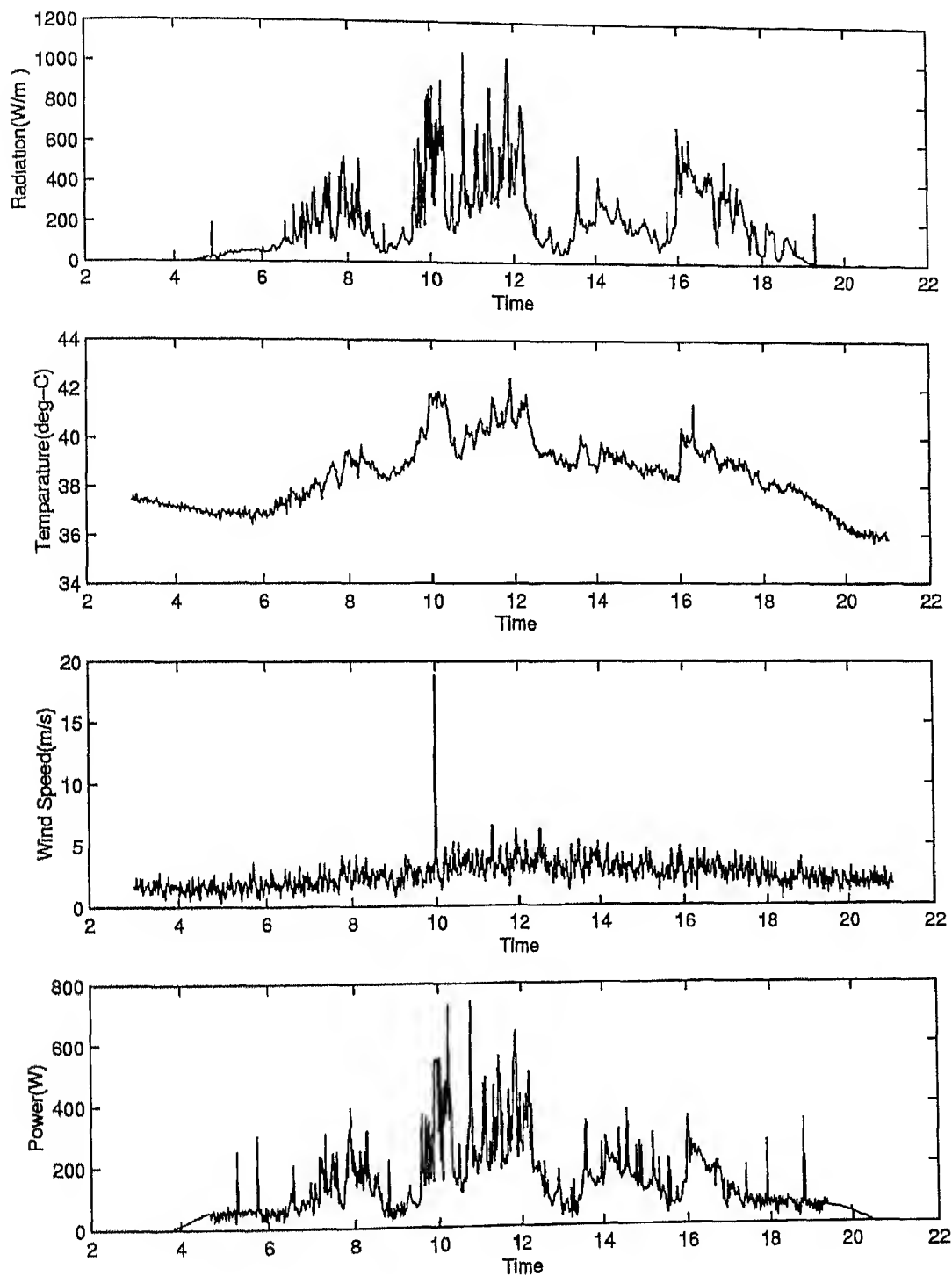
Figure A 5 Solar Data for June 20th 1994
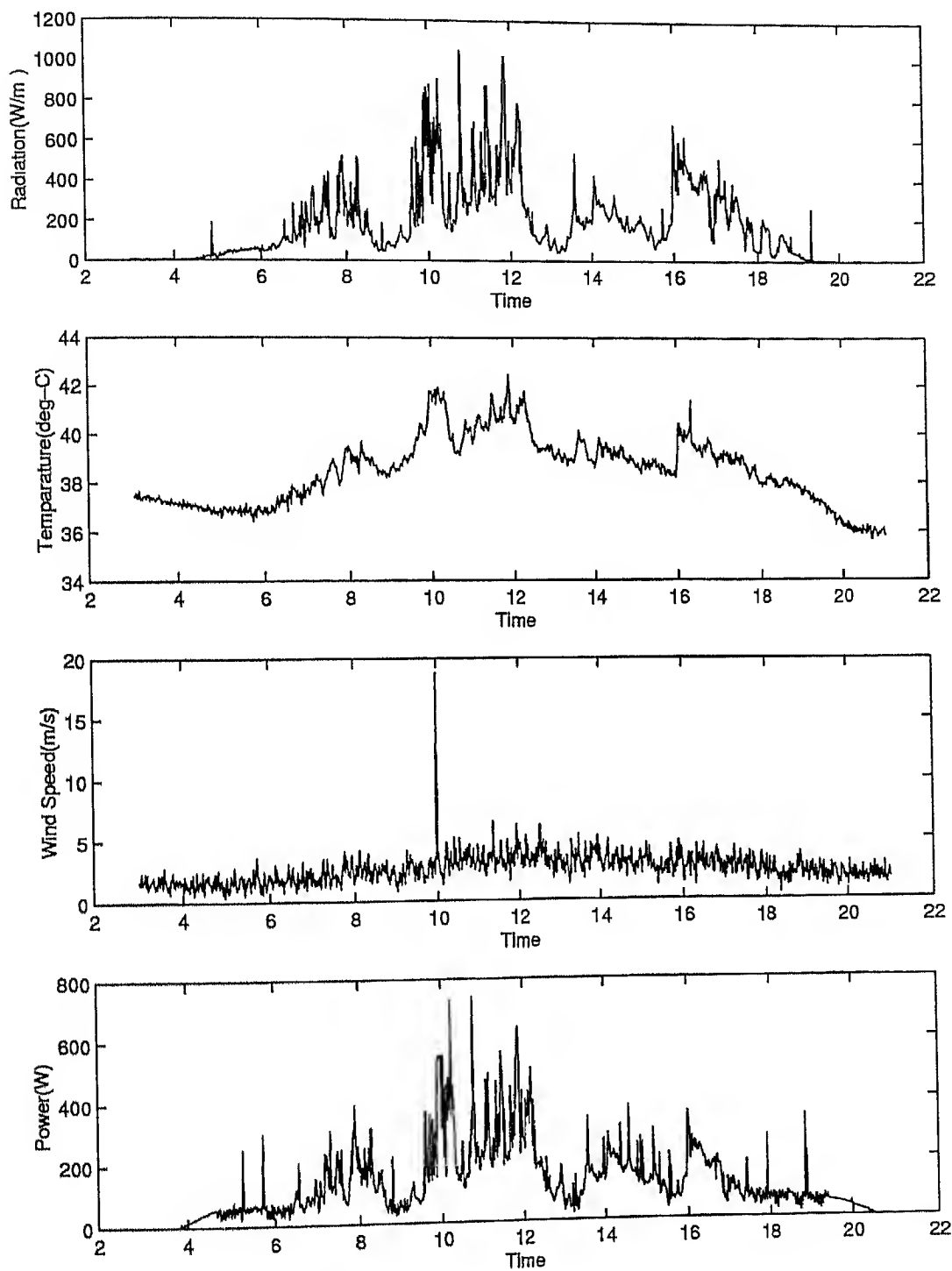
Figure A 6   Solar Data for June 21st 1994
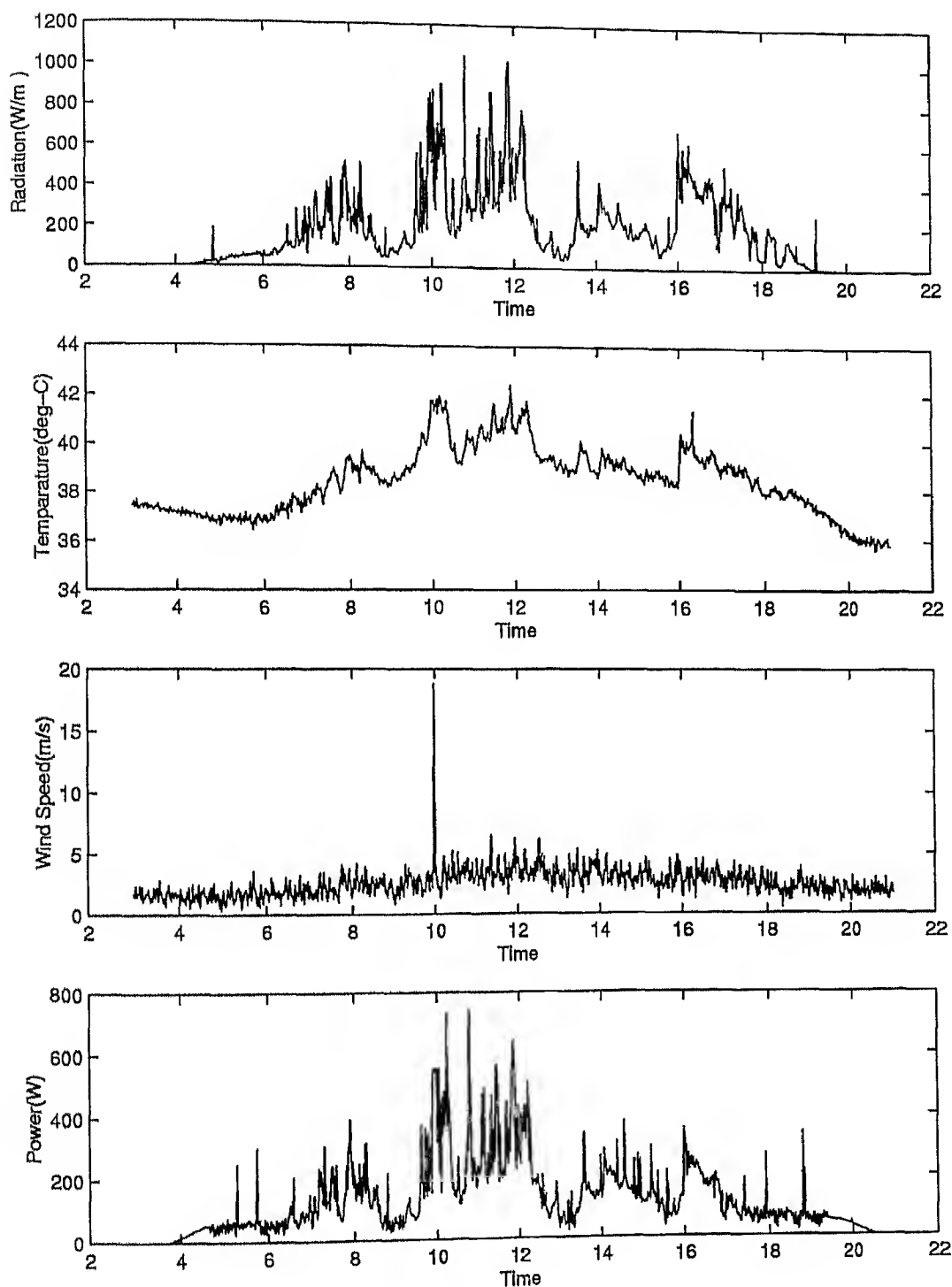
Figure A 7   Solar Data for June 22nd 1994
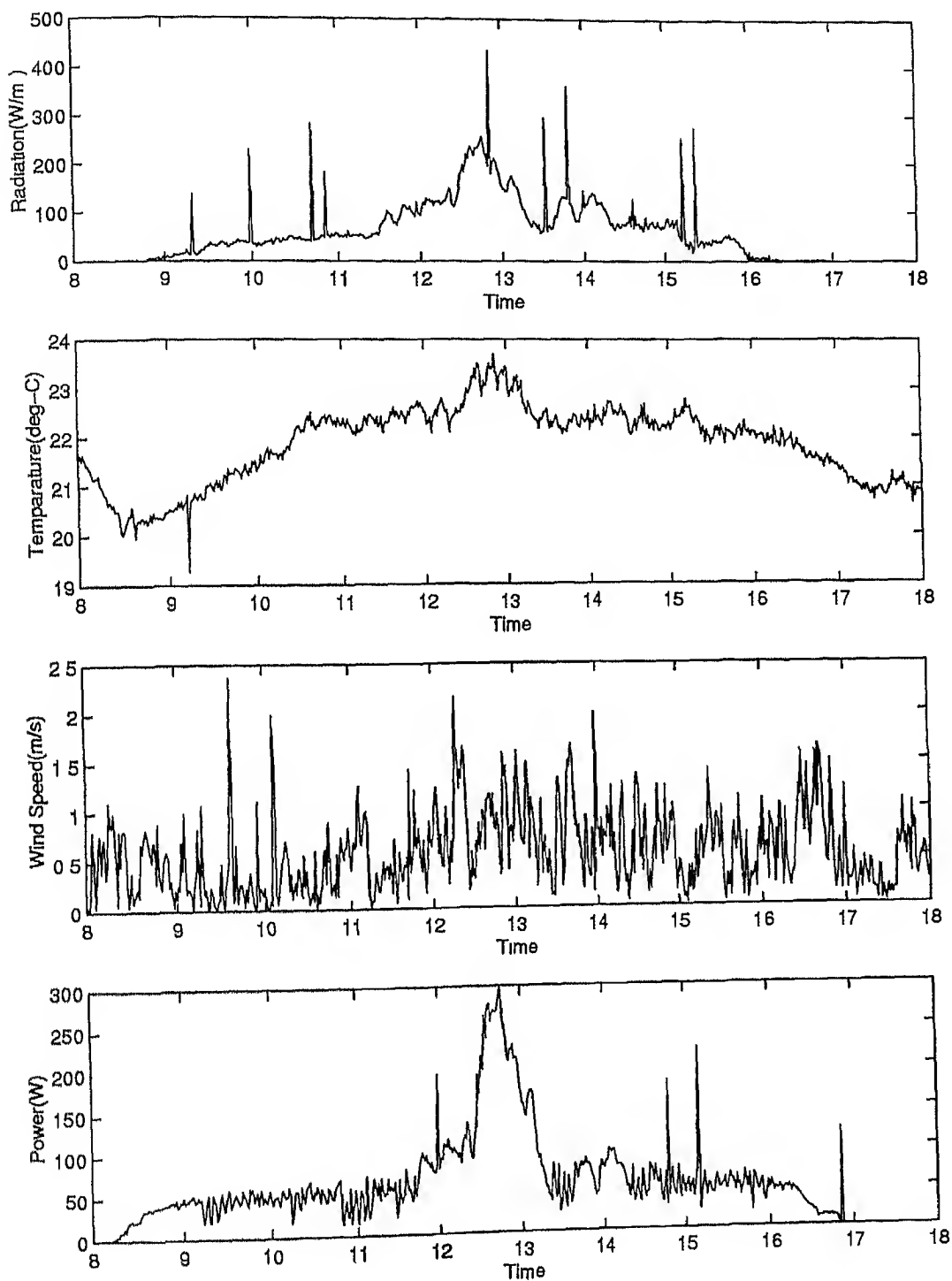
Figure A 8   Solar Data for January 9th 1994
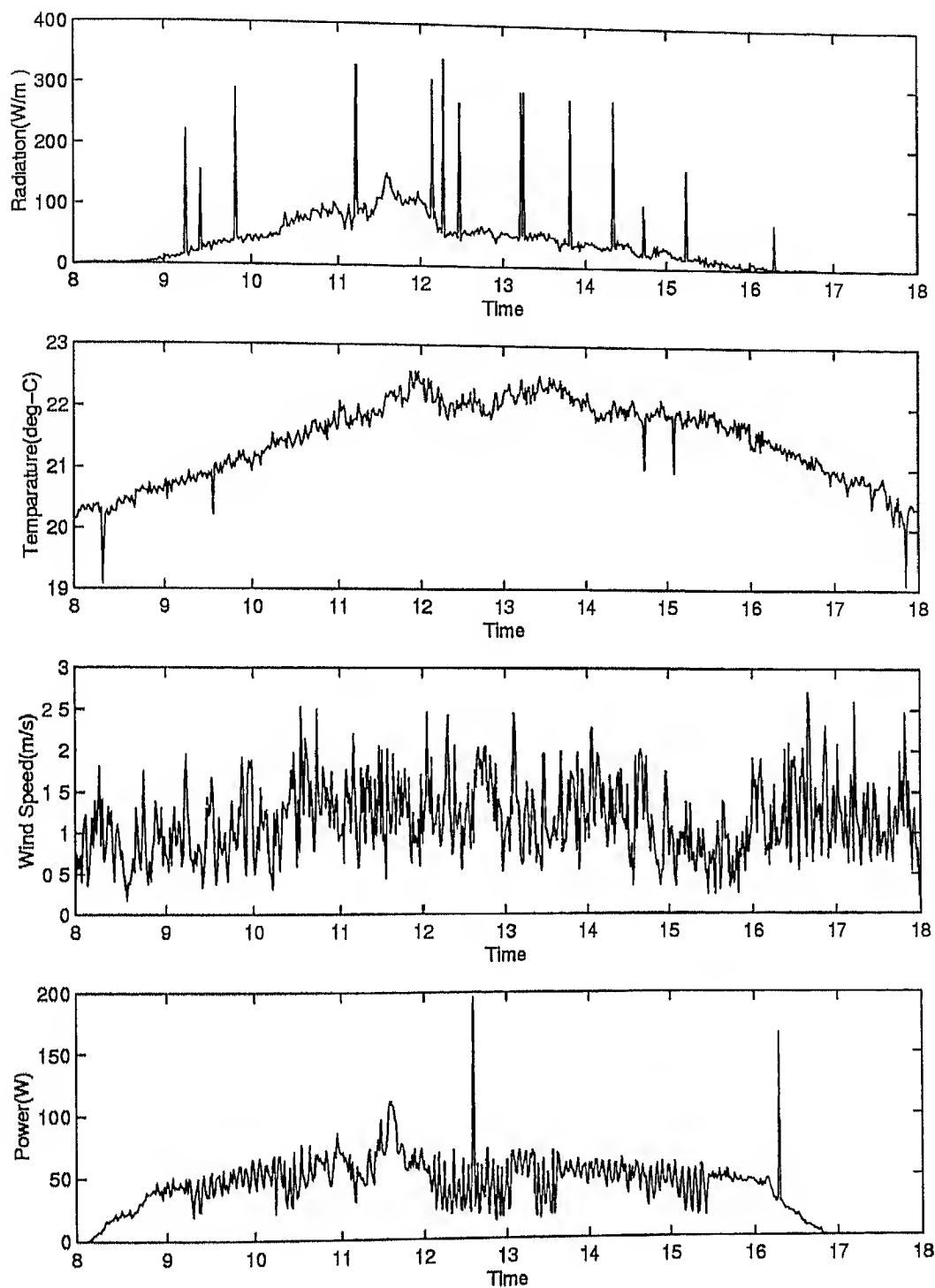
Figure A 9   Solar Data for January 10th 1994
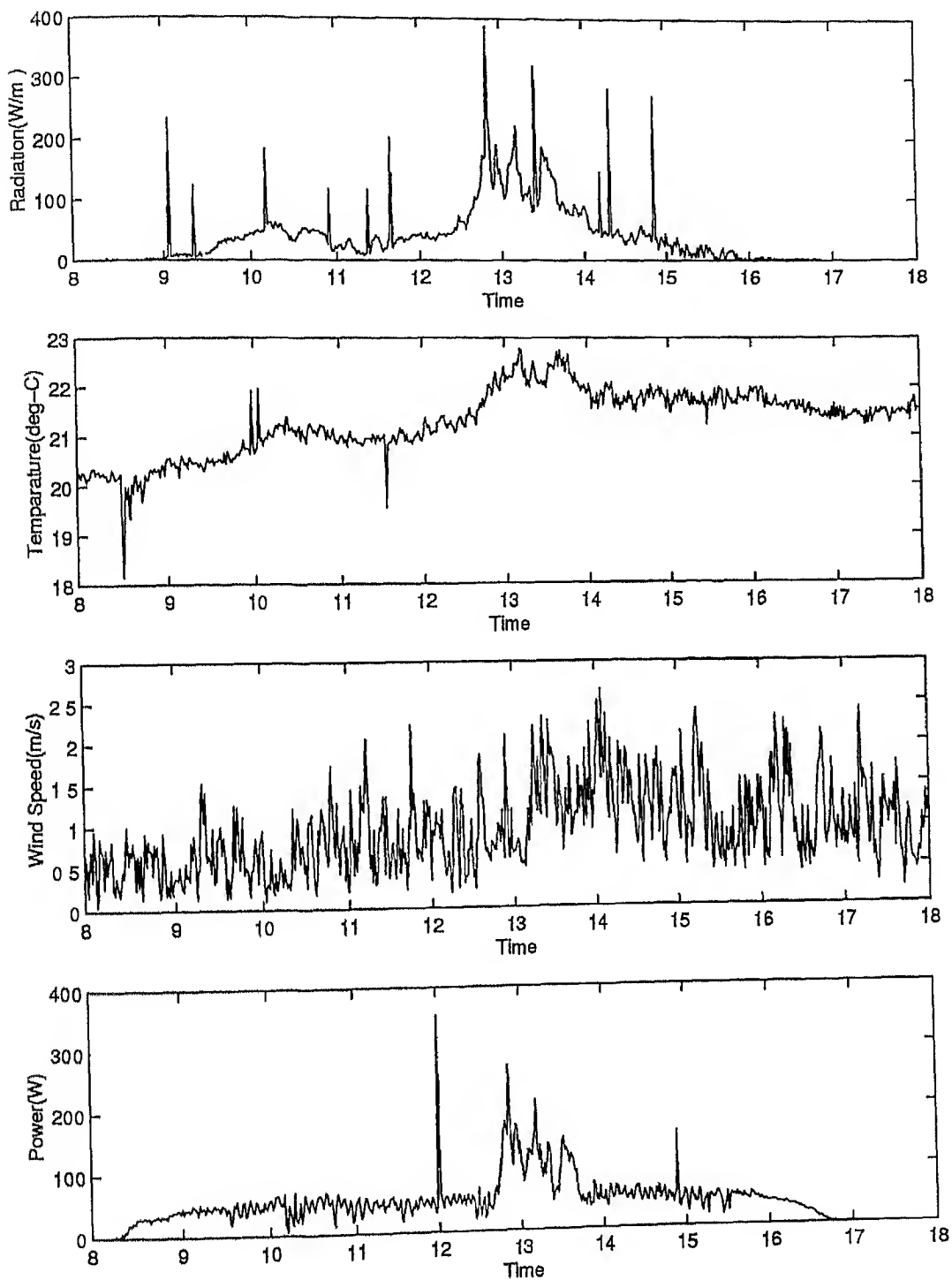
Figure A 10   Solar Data for January 11th 1994
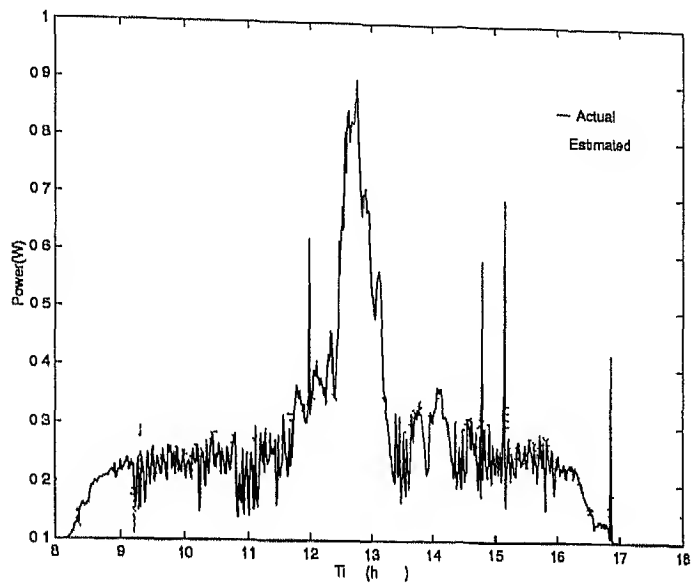
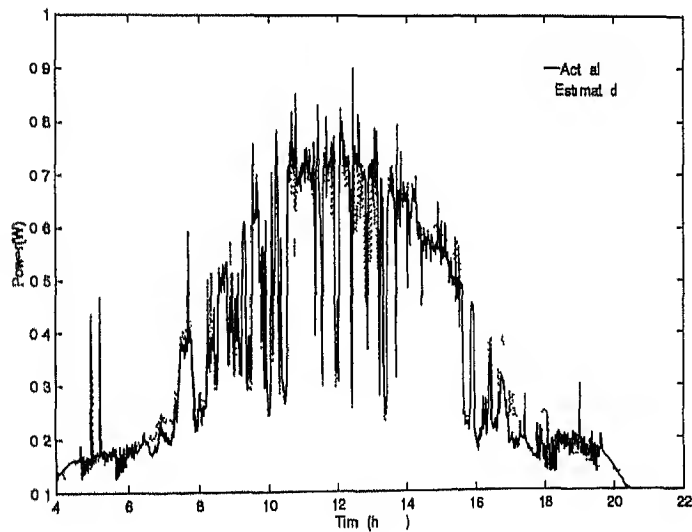Figure A 11   Phase I 9th january data neural network results

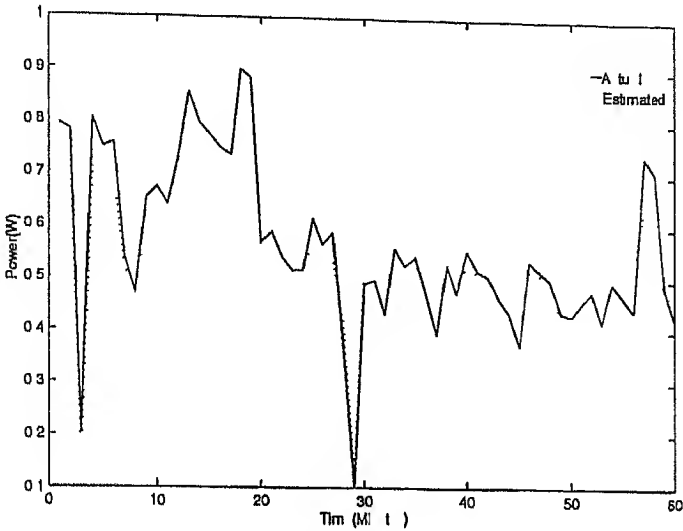Figure A 12   Phase I 20th june data neural network results

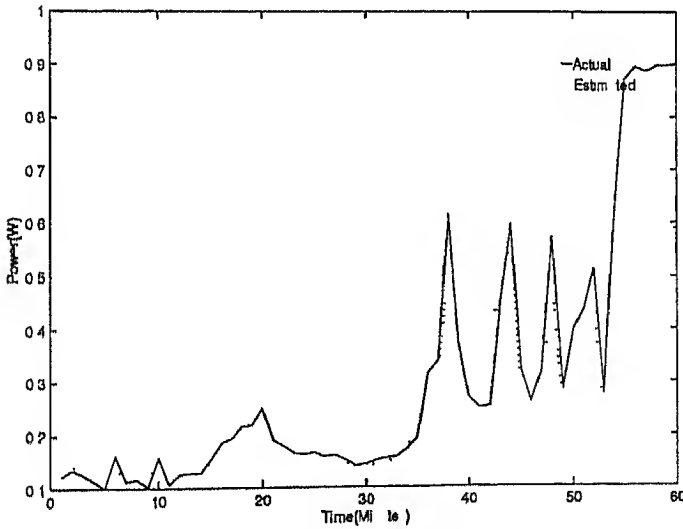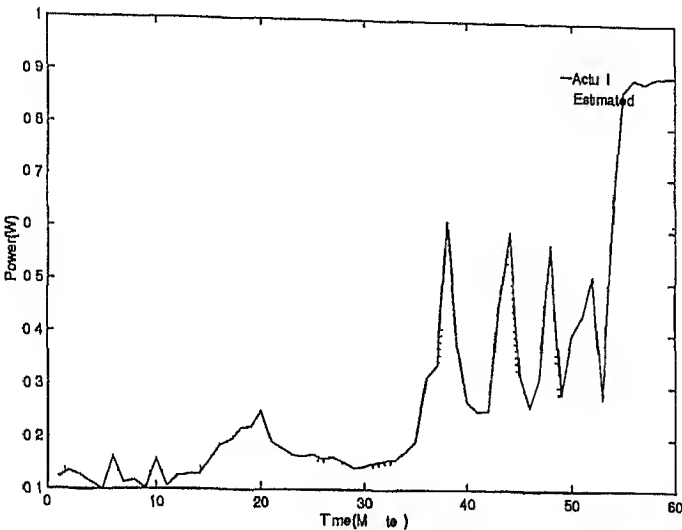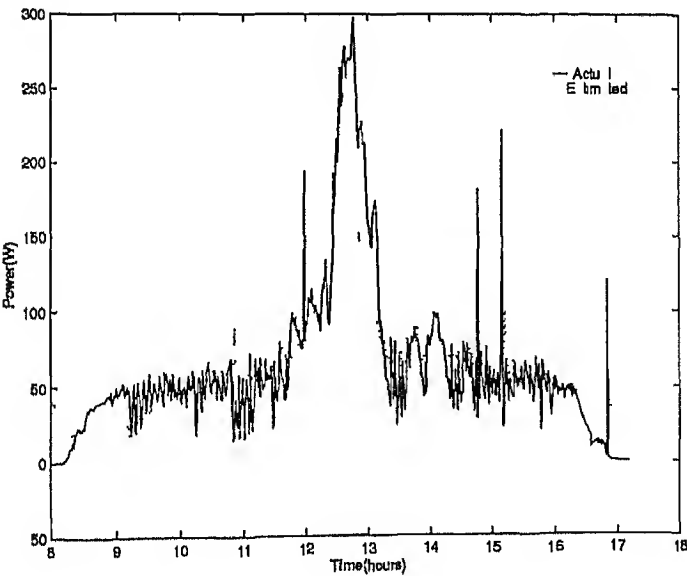Figure A 13   Phase I 20th june 14th hour data neural network results

Figure A 14   Phase I 22nd june 9th hour data neural network results

Figure A 15   Phase II 22nd june 9th hour data neural network results

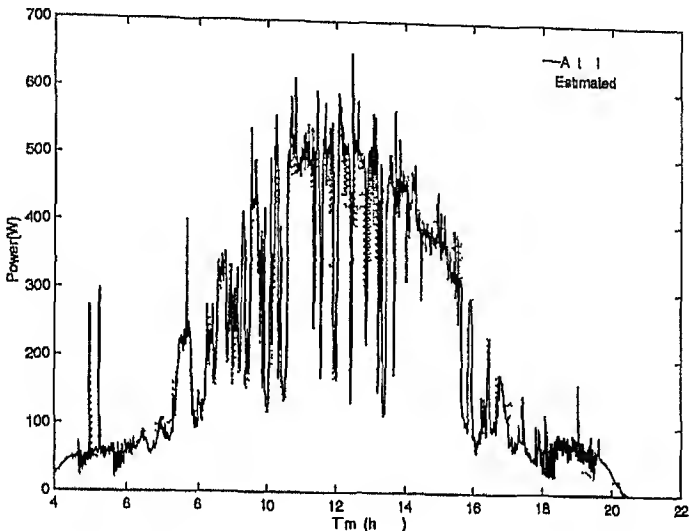Figure A 16   Phase I 9th january data multiple regression results

Figure A 17   Phase I 20th june data multiple regression results
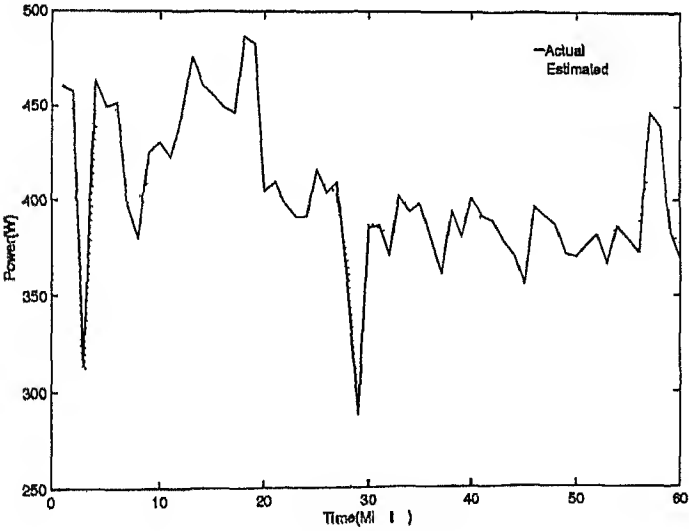

Figure A 18   Phase I 20th june 14th hour data multiple regression results
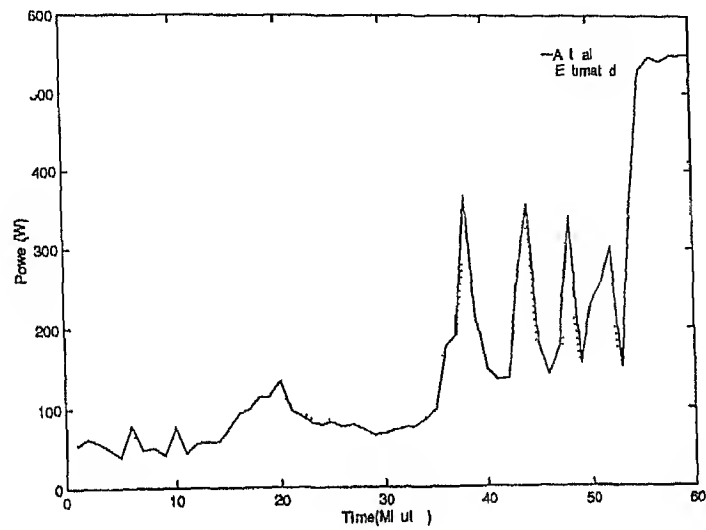
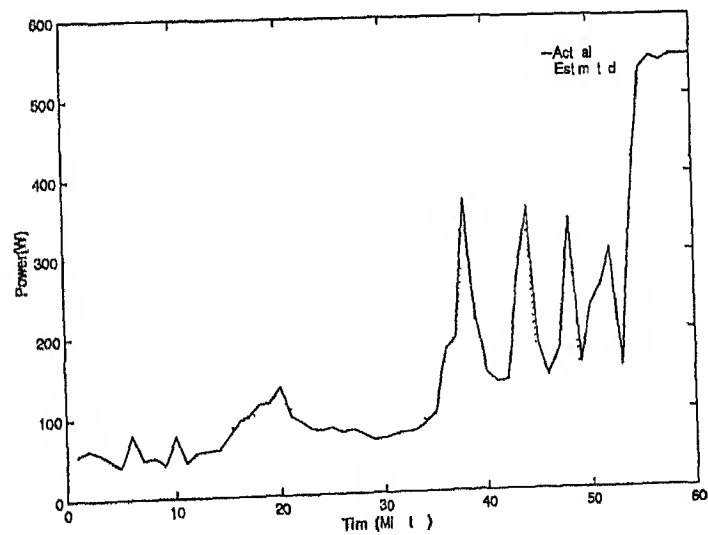Figure A 19   Phase I 22nd june 9th hour data multiple regression results

Figure A 20   Phase II 22nd june 9th hour data multiple regression results

Bibliography

[1] Alexiadis, M C , Dokopoulos, P S (IEEE member), Sahsamanoglou, H S , *Wind Speed and Power Forecasting based on Spatial Correlation Models*, IFFE Transactions on Energy Conversion, Vol 14, No 3, September 1999, pp 836 842

[2] Annan, R H , Herwing, L O *Photovoltaic industry and electric utility roles in the U S photovoltaic program*, Proceedings of the Biennial Congress of the International Solar Energy Society, Denver, p 331, 1991

[3] Applebaum, J *The quality of load matching in a direct coupling photovoltaic system*, IEEE Trans on Energy Conversion, Vol 2, No 4, pp 534 541, 1987

[4] Beyci, et a H G *Short term prediction of wind speed and power output of a wind turbine with neural networks*, Procd of the 5th European Wind Energy Conference, EWEC94, Thessaloniki, Greece, 1994, pp 349 352

[5] Bock, U , Nissen, J *Standardisiert Lastprofile fur Haushalte und Kleingewerbe*, Energiewirtschaftliche Tagesfragen, Jg 49 (1999), Heft 9, p606 610

[6] Bossanyi, E A *Stochastic wind prediction for wind turbine system control*, Procd of the 7th BWEA Wind Energy Conference, 27 29 March 1985, pp 219 226

[7] Chatfield, C *Statistics for Technology*, A couice in Applied Statistics,

[8] Edwards, A L *Multiple Regression and the analysis of variance and co variance*, W H Freeman & Co , pp 1-66, 1979

[9] Elsam *Wind power prediction tool in central dispatch centres*, Progress Report 3 of the LU project JOU2 CT92 0083, October 1993

[10] Hassoun, M H *Fundamentals of Artificial Neural Networks*, Prentice Hall of India private Limited New Delhi, 1998

[11] Haykin, S *Neural Networks*, Prentice Hall, Inc , New Jersey, 1999

[12] Hiyama, T , Kitabayashi, K *Neural Network Based Estimation of Maximum Power Generation from PV Module Using Environmental Information*, IEEE Transactions on Energy Conversion, Vol 12, No 3, September 1997

[13] Jewell, W T , Unruh, T D *Limits on cloud induced fluctuation in photovoltaic generation*, IEEE Transactions on Energy Conversion, Vol 5, No 1, March 1990

[14] Kaltschmitt, M , Wies(Hrsg), A *Erneuerbare Energien*, System technik, Wirtschaftlichkeit unweltaspekte ISBN3 540 59362 4 Springer verlag Berlin Heidelberg, 1995

[15] Kariniotakis, G , Nogaret, E , Stavrakakis, G *Wind power forecasting using advanced neural network models*, IEEE Transactions on Energy Conversion Vol 11 No 4 pp 762 767, December 1996

[16] Kosko, B *Neural Networks and Fuzzy Systems A Dynamic Systems Approach to Machine Intelligence*, Englewook Cliffs, NJ Prentice Hall, 1992

[17] Offer *Fifth renewable order for England and Wales*, Birmingham, September 1998

[18] Ponce de Leao, M T , Matos, M A *Fuzzy Models for Producers from Naturla Resources*, IECE proceding 1998

[19] Ross, T J *Fuzzy Logic with Engineering Applications* , McGraw Hill 1995

[20] Sakk, E , Thomas, R , Zimmerman, R *Power System Tournaments for a Deregulated Environment*, System Sciences, 1997, Proceedings of the Thirtieth Hawaii International Conference on System Sciences, p681-686

[21] Snoud, Z S , Jenkins, N , *Simple wind farm dynamic model* IEE Proc Generation Transimission Distribution, Vol 142, No 5 September 1995

[22] Schaeffer, G J , Boots, M G , Martens, J W Voogt, M H *Tradable Green Certificates A new market based incentive scheme for renewable energy Introduction and analysis*, Report from Netherlands Energy Research Foundation ECN, March 1999

[23] Schwartz, R J *Photovoltaic Power Generation*, Procedings of the IEEE, Vol 81, No 3, March 1993

[24] Shuhui, Li, Don, C , Wunsch , O'Hair, E , Giesselmann, M G *Neural Network for Wind Power Generation with Compressing Function,*

[25] Twidell, J *Renewable Energy Implementation and benefits* IEE 2nd International Conference on Advances in Power System Control, Operation and Management, December 1993, Hong Kong

[26] Wilson, R E , Spera, D A *Aerodynamic Behavior of Wind Turbines*, in Wind Turbine Technology, AMSE Press, chap 5, pp 215 282, 1995

[27] Winter, C J , Sizmann, R L , Vant Hull, L L *Solar Power Plants*, Fundamentals, Technology Systems and Economics Springer Verlag Berlin Heidelburg 1991

[28] Zurada, J M *Introduction to artificial neural systems* St Paul[u a] West Publishing, 1992

A 130924

A 130924